

OM protein - protein search, using sw model

Run on: March 1, 2005, 08:40:57 ; Search time 205.443 Seconds
(without alignments)
1690.542 Million cell updates/sec

Title: US-10-624-932-2
Perfect score: 4791
Sequence: 1 MAVRPGLWPALLGIVLAWL.....AVAGLGQPDAGLFTVSEAE 898

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	4791	100.0	898	5	AAU85403	Aau85403 Human pro
2	4791	100.0	898	8	ADH71618	Adh71618 Human pro
3	4781	99.8	898	5	AAU97899	Aau97899 Human net
4	4698.5	98.1	899	5	AAU79939	Aau79939 Human UNC
5	4698.5	98.1	899	7	ADG42569	Adg42569 Novel hum
6	4698.5	98.1	899	8	ADH71610	Adh71610 Human pro
7	4698.5	98.1	899	8	ADH71628	Adh71628 Human pro
8	4697.5	98.0	899	8	ADH71636	Adh71636 Human pro
9	4695.5	98.0	899	8	ADH71642	Adh71642 Human pro

10	4695.5	98.0	899	8	ADH71640	Adh71640	Human	pro
11	4694.5	98.0	899	8	ADH71632	Adh71632	Human	pro
12	4694.5	98.0	899	8	ADH71638	Adh71638	Human	pro
13	4693.5	98.0	899	8	ADH71634	Adh71634	Human	pro
14	4691.5	97.9	899	8	ADH71648	Adh71648	Human	pro
15	4691.5	97.9	899	8	ADH71630	Adh71630	Human	pro
16	4691.5	97.9	899	8	ADH71644	Adh71644	Human	pro
17	4686.5	97.8	899	8	ADH71646	Adh71646	Human	pro
18	4684.5	97.8	899	8	ADH71650	Adh71650	Human	pro
19	4683	97.7	898	8	ADH71626	Adh71626	Human	pro
20	4638	96.8	898	2	AAW78898	Aaw78898	Rat	UNC-5
21	4638	96.8	898	5	AAU10543	Aau10543	Rat	netri
22	4638	96.8	898	5	AAU97900	Aau97900	Rat	netri
23	4620	96.4	898	7	ADG42580	Adg42580	Rat	trans
24	4526.5	94.5	943	4	AAM79128	Aam79128	Human	pro
25	4413	92.1	842	5	AAU74818	Aau74818	Human	REP
26	4413	92.1	842	8	ADL06574	Adl06574	Human	tum
27	4360	91.0	833	8	ADH71622	Adh71622	Human	pro
28	4275	89.2	817	8	ADH71624	Adh71624	Human	pro
29	2839	59.3	544	7	ADG42581	Adg42581	Human	tra
30	2815	58.8	556	2	AAW78899	Aaw78899	Human	UNC
31	2768	57.8	931	7	ADG42582	Adg42582	Mouse	tra
32	2755	57.5	931	4	AAB50691	Aab50691	Human	UNC
33	2755	57.5	931	7	ADE63098	Ade63098	Human	Pro
34	2755	57.5	931	7	ABU64297	Abu64297	Human	thr
35	2755	57.5	931	8	ADR99258	Adr99258	Human	unc
36	2755	57.5	982	4	ABG11551	Abg11551	Novel	hum
37	2741	57.2	931	7	ADG42584	Adg42584	Human	tra
38	2736	57.1	964	8	ADR99250	Adr99250	Human	lRO
39	2696	56.3	929	7	ADG42583	Adg42583	Human	tra
40	2578.5	53.8	945	7	ADE63096	Ade63096	Rat	Prote
41	2571.5	53.7	943	2	AAW78900	Aaw78900	Rat	UNC-5
42	2563.5	53.5	933	5	AAO18734	Aao18734	Human	NOV
43	2563.5	53.5	933	5	AAO18735	Aao18735	Human	NOV
44	2558.5	53.4	945	4	AAU12244	Aau12244	Human	PRO
45	2558.5	53.4	945	6	ABO17688	Abo17688	Novel	hum

ALIGNMENTS

RESULT 1

AAU85403

ID AAU85403 standard; protein; 898 AA.

XX

AC AAU85403;

XX

DT 21-MAY-2002 (first entry)

XX

DE Human protein NOV1.

XX

KW Human; NOX; cardiomyopathy; atherosclerosis; diabetes;

KW cell signal processing disorder; metabolic disorder; obesity; infection;

KW anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;

KW Alzheimer's disease; Parkinson's disease; immune disorder;

KW haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;

KW osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;

KW myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
 KW psychosis; neurological disorder; anxiety; schizophrenia;
 KW manic depression; dementia; dyskinesia; Huntington's disease;
 KW Gilles de la Tourette's syndrome; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200210216-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 30-JUL-2001; 2001WO-US024225.
 XX
 PR 28-JUL-2000; 2000US-0221409P.
 PR 04-AUG-2000; 2000US-0222840P.
 PR 04-AUG-2000; 2000US-0223752P.
 PR 04-AUG-2000; 2000US-0223762P.
 PR 04-AUG-2000; 2000US-0223769P.
 PR 04-AUG-2000; 2000US-0223770P.
 PR 14-AUG-2000; 2000US-0225146P.
 PR 15-AUG-2000; 2000US-0225392P.
 PR 15-AUG-2000; 2000US-0225470P.
 PR 16-AUG-2000; 2000US-0225697P.
 PR 01-FEB-2001; 2001US-0263662P.
 PR 05-APR-2001; 2001US-0281645P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Padigar M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM;
 PI Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;
 XX
 DR WPI; 2002-180074/23.
 DR N-PSDB; ABK37922.
 XX
 PT New isolated cytoplasmic, nuclear, membrane bound, or secreted
 PT polypeptide, useful for treating cardiomyopathy, atherosclerosis,
 PT infections, cancer, neurodegenerative, metabolic, hematopoietic and
 PT immune disorders.
 XX
 PS Claim 1; Page 11; 213pp; English.
 XX
 CC The invention relates to an isolated cytoplasmic, nuclear, membrane
 CC bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature
 CC form. Also included are the nucleic acids encoding the NOVX proteins, a
 CC vector comprising the nucleic acid, a cell comprising the vector, an anti
 CC -NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the
 CC antibody are useful for treating or preventing a NOVX-associated
 CC disorder, where the disorder is selected from cardiomyopathy,
 CC atherosclerosis, diabetes, a disorder related to cell signal processing
 CC and metabolic pathway modulation, metabolic disorders, obesity,
 CC infectious disease, anorexia, cancer-associated cachexia, cancer,
 CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
 CC immune disorders, haematopoietic disorders, and the various
 CC dyslipidaemias, metabolic disturbances associated with obesity, the
 CC metabolic syndrome X and wasting disorders associated with chronic
 CC diseases, bacterial, fungal, protozoal and viral infections, pain,
 CC bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's

CC disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina
 CC pectoris, myocardial infarction, ulcer, allergy, benign prostatic
 CC hypertrophy, and psychotic and neurological disorders, including anxiety,
 CC schizophrenia, manic depression, delirium, dementia, and dyskinesias,
 CC such as Huntington's disease and Gilles de la Tourette's syndrome. The
 CC nucleic acid is useful in gene therapy. The present sequence represents a
 CC NOVX protein

XX

SQ Sequence 898 AA;

Query Match 100.0%; Score 4791; DB 5; Length 898;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 898; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Db	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Qy	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL	120
Db	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL	120
Qy	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAE	180
Db	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAE	180
Qy	181	VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY	240
Db	181	VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Qy	301	WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVA	360
Db	301	WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVA	360
Qy	361	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHLL	420
Db	361	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHLL	420
Qy	421	TIQPDLS TTTT TYQGS LCPQDGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSEAE EFVS	480
Db	421	TIQPDLS TTTT TYQGS LCPQDGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSEAE EFVS	480
Qy	481	RLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLHK	540
Db	481	RLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLHK	540
Qy	541	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPDSWSLRLKKQSCEGSW	600
Db	541	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPDSWSLRLKKQSCEGSW	600
Qy	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACT	660
Db	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACT	660

Qy 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720
 |||
 Db 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720

Qy 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSFSINF 780
 |||
 Db 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSFSINF 780

Qy 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRGGADWRTLAQKL 840
 |||
 Db 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRGGADWRTLAQKL 840

Qy 841 HLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898
 |||
 Db 841 HLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898

RESULT 2

ADH71618

ID ADH71618 standard; protein; 898 AA.

XX

AC ADH71618;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21e SEQ ID NO:514.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
 KW obesity; diabetes; infectious disease; metabolic syndrome X;
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.

Db	121	 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAAE	180
Qy	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY	240
Db	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Qy	301	WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDLDTNRNCTSDLCVHSASGPEDVA	360
Db	301	WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDLDTNRNCTSDLCVHSASGPEDVA	360
Qy	361	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHLL	420
Db	361	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHLL	420
Qy	421	TIQPDLS TTTT TYQGS LCPQDGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSEAEFEVS	480
Db	421	TIQPDLS TTTT TYQGS LCPQDGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSEAEFEVS	480
Qy	481	RLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHK	540
Db	481	RLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHK	540
Qy	541	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPDWSLRLKKQSCGWS	600
Db	541	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPDWSLRLKKQSCGWS	600
Qy	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACT	660
Db	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACT	660
Qy	661	SLEYNIRVYCLHDT HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW	720
Db	661	SLEYNIRVYCLHDT HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW	720
Qy	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSFSINF	780
Db	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSFSINF	780
Qy	781	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL	840
Db	781	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL	840
Qy	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	898
Db	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	898

RESULT 3

AAU97899

ID AAU97899 standard; protein; 898 AA.

XX

AC AAU97899;

XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Human netrin binding membrane receptor UNC5H-1 protein.
 XX
 KW Netrin binding membrane receptor; receptor; UNC5H-1; human; nootropic;
 KW neuroprotective; cytostatic; antiparkinsonian; cerebroprotective; cancer;
 KW central nervous system; CNS; stroke; Parkinson's disease;
 KW multiple sclerosis; Alzheimer's disease.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 152. .223
 FT /note= "Immunoglobulin domain "
 FT Domain 247. .294
 FT /note= "Thrombospondine type 1 domain "
 FT Domain 302. .348
 FT /note= "Thrombospondine type 1 domain"
 FT Region 361. .382
 FT /note= "Transmembrane region"
 FT Domain 495. .598
 FT /note= "ZU5 domain"
 FT Domain 817. .897
 FT /note= "Death domain"
 XX
 PN WO200233080-A2.
 XX
 PD 25-APR-2002.
 XX
 PF 15-OCT-2001; 2001WO-EP011891.
 XX
 PR 16-OCT-2000; 2000US-0240061P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Koehler RH;
 XX
 DR WPI; 2002-463314/49.
 DR N-PSDB; ABK52891.
 XX
 PT Novel human netrin binding membrane receptor polypeptide and
 PT polynucleotides for identifying modulating agents useful in treating
 PT diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
 PT Alzheimer's disease.
 XX
 PS Claim 1; Fig 2; 94pp; English.
 XX
 CC This invention relates to the DNA and protein sequences of a novel
 CC purified human netrin binding membrane receptor, UNC5H-1. The DNA
 CC sequence of the invention is useful as a probe for detecting a nucleic
 CC acid encoding the UNC5H-1 protein in a biological sample. The sequences
 CC of the invention are useful to screen for agents which decrease the
 CC activity of the UNC5H-1 protein. The sequences are also useful for
 CC screening agents which regulate (modulate) the activity of the protein of
 CC the invention. A pharmaceutical composition containing the protein of the
 CC invention or a reagent that modulates the activity of the UNC5H-1 protein

CC may be useful for treating a UNC5H-1 dysfunction related disease such as
CC cancer or a central nervous system (CNS) disorders (e.g, Parkinson's
CC disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
CC proteins comprising the UNC5H-1 protein are useful for generating
CC antibodies and for in various assay systems, and the protein can be used
CC as a bait protein in a two-hybrid assay or three-hybrid assay. The method
CC of the invention is useful for detecting a coding sequence for the UNC5H-
CC 1 protein. The present sequence represents the human netrin binding
CC membrane receptor UNC5H-1 protein of the invention

XX

SQ Sequence 898 AA;

Query Match 99.8%; Score 4781; DB 5; Length 898;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 896; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
      |||
Db      1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLE 120
      |||
Db     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLE 120

Qy    121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE 180
      |||
Db    121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE 180

Qy    181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY 240
      |||
Db    181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY 240

Qy    241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGS 300
      |||
Db    241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGS 300

Qy    301 WSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRNCTSDLCVHSASGPEDVA 360
      |||
Db    301 WSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRNCTSDLCVHTASGPEDVA 360

Qy    361 LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSQFQVSIKPSKADNPHELL 420
      |||
Db    361 LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSQFQVSIKPSKADNPHELL 420

Qy    421 TIQPDLSSTTTTYYQGSCLCPQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFEVS 480
      |||
Db    421 TIQPDLSSTTTTYYQGSCLCPQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFEVS 480

Qy    481 RLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHK 540
      |||
Db    481 RLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHK 540

Qy    541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSW 600
      |||
Db    541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSW 600

Qy    601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLL FAPVACT 660
```

Db	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA ^{AA} AKRLKLLLFAPVACT	660
Qy	661	SLEYNIRVYCLHDT ^{HD} DALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLR ^{LS} SIH ^D VPSSLW	720
Db	661	SLEYNIRVYCLHDT ^{HD} DALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLR ^{LS} SIH ^D VPSSLW	720
Qy	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSFSINF	780
Db	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSFSINF	780
Qy	781	NITKDTRFAELLA ^{LE} SEAGVPALVGPSAFKIPFLIRQKI ^{ISS} LDPPCRRGADWRTLAQKL	840
Db	781	NITKDTRFAELLA ^{LE} SEAGVPALVGPSAFKIPFLIRQKI ^{ISS} LDPPCRRGADWRTLAQKL	840
Qy	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE ^C	898
Db	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE ^C	898

RESULT 4

AAU79939

ID AAU79939 standard; protein; 899 AA.

XX

AC AAU79939;

XX

DT 15-JUL-2002 (first entry)

XX

DE Human UNC5-like protein NOV1.

XX

KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
 KW cell signal processing; metabolic pathway modulation; cancerous tissue;
 KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
 KW chromosome 13.

XX

OS Homo sapiens.

XX

PN WO200229038-A2.

XX

PD 11-APR-2002.

XX

PF 04-OCT-2001; 2001WO-US031377.

XX

PR 04-OCT-2000; 2000US-0237862P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Herrmann JL, Rastelli L, Shimkets RA;

XX

DR WPI; 2002-340104/37.

DR

N-PSDB; ABK49422.

XX

PT Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for

PT

treating cardiomyopathy, artherosclerosis, and cancer.

XX

PS Claim 1; Page 9; 180pp; English.

XX

Db 540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSSWSLRLKKQSCEGS 599
 Qy 600 WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVA 658
 || ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVA 659
 Qy 659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPVSS 718
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPVSS 719
 Qy 719 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 778
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
 Qy 779 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 838
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
 Qy 839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 899

RESULT 5

ADG42569

ID ADG42569 standard; protein; 899 AA.

XX

AC ADG42569;

XX

DT 26-FEB-2004 (first entry)

XX

DE Novel human NOV1.

XX

KW cytostatic; gene therapy; NOVX-agonist; NOVX-antagonist; pharmaceutical;

KW NOVX-associated disorder; cancer; human.

XX

OS Homo sapiens.

XX

PN US2003204052-A1.

XX

PD 30-OCT-2003.

XX

PF 04-OCT-2001; 2001US-00970944.

XX

PR 04-OCT-2000; 2000US-0237862P.

XX

PA (HERR/) HERRMANN J L.

PA (RAST/) RASTELLI L.

PA (SHIM/) SHIMKETS R A.

XX

PI Herrmann JL, Rastelli L, Shimkets RA;

XX

DR WPI; 2003-900673/82.

DR N-PSDB; ADG42568.

XX

PT New NOVX gene or NOVX-specific antibody, useful for preparing a

PT composition for treating or preventing a NOVX-associated disorder, e.g.,

PT cancer.

XX

PS Claim 1; SEQ ID NO 2; 118pp; English.

XX

CC The invention describes a new isolated polypeptide comprising: a
CC polypeptide or its mature form comprising a sequence not given in the
CC specification; or a variant of (A), where one or more amino acid residues
CC in the variant differs in no more than 15% from the amino acid sequence
CC of the mature form. The pharmaceutical composition may be administered
CC via oral, transdermal, rectal or parenteral route. The polypeptide,
CC nucleic acid or antibody is useful for preparing a composition for
CC treating or preventing a NOVX-associated disorder, e.g., cancer. This is
CC the amino acid sequence of a human NOVX protein.

XX

SQ Sequence 899 AA;

Query Match 98.1%; Score 4698.5; DB 7; Length 899;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

Qy	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Db	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Qy	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL	120
Db	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRRQVEKVFGL	120
Qy	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAE	180
Db	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAE	180
Qy	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY	240
Db	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNV-QKTACATLCPVDG	299
		: :	
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVHDRTVSSLLVSVDG	300
Qy	300	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	359
Db	301	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	360
Qy	360	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSQFQPVSIKPSKADNPHL	419
Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSQFQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLSTTTTTYQGSICPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFEV	479
Db	421	LTIQPDLSTTTTTYQGSICPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFEV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTNFGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTNFGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRLKKQSCEGS	599

Db	540		KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPDSWSLRLKKQSCEGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA	AAKRLKLLLFAPVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA	AAKRLKLLLFAPVA	659
Qy	659	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLR	LSIHDPSS	718
Db	660	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLR	LSIHDPSS	719
Qy	719	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI		778
Db	720	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI		779
Qy	779	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ		838
Db	780	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ		839
Qy	839	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	C	898
Db	840	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	C	899

RESULT 6

ADH71610

ID ADH71610 standard; protein; 899 AA.

XX

AC ADH71610;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21a SEQ ID NO:506.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.

Qy	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL	120
Db	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGSEPTMEVRINVSRRQVEKVFGL	120
Qy	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAE	180
Db	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAE	180
Qy	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY	240
Db	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNV-QKTACATLCPVDG	299
		: :	
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVHVRTVSSLLVSDG	300
Qy	300	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	359
Db	301	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	360
Qy	360	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL	419
Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLS-TTTTQYQSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Db	421	LTIQPDLS-TTTTQYQSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Qy	480	SRLSTQNYFRSLPRGTSNMITYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMITYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGS	599
Db	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLLFAPVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLLFAPVA	659
Qy	659	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSS	718
Db	660	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSS	719
Qy	719	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI	778
Db	720	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI	779
Qy	779	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	838
Db	780	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	839
Qy	839	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	898
Db	840	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	899

RESULT 7

ADH71628

ID ADH71628 standard; protein; 899 AA.

XX

AC ADH71628;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21j SEQ ID NO:524.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.

PR 12-JUN-2002; 2002US-0387934P.

PR 12-JUN-2002; 2002US-0387960P.

PR 12-JUN-2002; 2002US-0388022P.

PR 12-JUN-2002; 2002US-0388096P.

PR 13-JUN-2002; 2002US-0389123P.

PR 14-JUN-2002; 2002US-0389118P.

PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR N-PSDB; ADH71627.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 524; 1880pp; English.

XX

The invention relates to a novel isolated polypeptide (NOVX). A polypeptide of the invention has cytostatic, immunomodulator, neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and antilipaemic activity, and may have a use in gene therapy, and as a vaccine. The polypeptides are encoded by NOVX polynucleotides comprising any of the 303 fully defined nucleotide sequences given in the specification. The polypeptide is useful in the manufacture of a medicament for treating a syndrome associated with a human disease. The polypeptide, polynucleotide and antibody are useful in diagnosing, treating or preventing NOVX-associated disorders, e.g. cancer, cachexia, Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are further used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine, and pharmacogenomics. The present sequence represents a NOVX polypeptide of the invention.

XX

SO Sequence 899 AA;

Query Match 98.1%; Score 4698.5; DB 8; Length 899;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

[illegible]

Db 1 MAVRPGLWPALLGIVLAAWLRGSGAOOSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

[illegible]

Db 61 VLLVCKAVPATQIFFKCNGEWVROVDHVIERSTDGSSGEPTMEVRINVSROOVEKVFGL 120

Qy 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAE 180
| | | | |

Db 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEOEPLAKEVSLEOGIVLPCRPPPEGIPPAE 180

[illegible]

Db 181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240

QY 241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPLNGGAFCEGQNV-QKTACATLCVPVDG 299
| | | | | | | | | | | | | | | | | | | | : | : | | |

Db 241 VNGGWSTWTEWSVCSASCGRGWOKRSRSC TNPAPLNGGAFCEGONVHDRTVSSLLVSVDG 300

QY 300 SWSPWSKWSACGLDCTHWRSRECS DPAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDV 359
| | | | |

Db 301 SWSPWSKWSACGLDCTHWRSRECS DPAPRNGGEECOGTDLDTRNCTSDLCVHSASGPEDV 360

QY 360 ALYVGLIAVAVCLVL~~LLLL~~VLIILVYCRKKEGLDSDVADSSILTSGFQPVSIIKPSKADNPHL 419
 |||||

Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSQFQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Db	421	LTIQPDLS-TTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSSWRLKKQSCGS	599
Db	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSSWRLKKQSCGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLLFAPVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLLFAPVA	659
Qy	659	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPVSS	718
Db	660	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPVSS	719
Qy	719	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSFSI	778
Db	720	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSFSI	779
Qy	779	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQ	838
Db	780	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQ	839
Qy	839	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAE	898
Db	840	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAE	899

RESULT 8

ADH71636

ID ADH71636 standard; protein; 899 AA.

XX

AC ADH71636;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21n SEQ ID NO:532.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
 KW obesity; diabetes; infectious disease; metabolic syndrome X;
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.

PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR N-PSDB; ADH71635.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 532; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence

CC represents a NOVX polypeptide of the invention.

XX

SQ Sequence 899 AA;

Query Match 98.0%; Score 4697.5; DB 8; Length 899;

Best Local Similarity 98.6%; Pred. No. 0;

Matches 887; Conservative 3; Mismatches 7; Indels 3; Gaps 3;

Qy	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Db	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Qy	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGL	120
Db	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSQQVEKVFGL	120
Qy	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE	180
Db	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE	180
Qy	181	VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASA	240
Db	181	VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASA	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNV-QKTACATLCPVDG	299
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVHDTVSSLLVSDG	300
Qy	300	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	359
Db	301	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	360
Qy	360	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHL	419
Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLSTTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Db	421	LTIQPDLS-TTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTNFNLGGRIMIPNTGISLLIPPDAIPRGIYEIYTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTNFNLGGRIMIPNTGISLLIPPDAIPRGIYEIYTLH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGS	599
Db	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYIFTEQLGRFALVGEALSVA	659
Qy	659	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS	718
Db	660	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS	719
Qy	719	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWWVQVEGDGQSFSI	778

Db	720	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI	779
Qy	779	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLQ	838
Db	780	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLQ	839
Qy	839	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	898
Db	840	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	899

RESULT 9

ADH71642

ID ADH71642 standard; protein; 899 AA.

XX

AC ADH71642;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21q SEQ ID NO:538.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
 KW obesity; diabetes; infectious disease; metabolic syndrome X;
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.

Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNV-QKTACATLCPVDG	299
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVHVRTVSSLLVSDG	300
Qy	300	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDV	359
Db	301	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDV	360
Qy	360	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSGFQPVSIKPSKADNPHL	419
Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSGFQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Db	421	LTIQPDLS-TTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGS	599
Db	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAALKRLKLLFAPVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAALKRLKLLFAPVA	659
Qy	659	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPVPS	718
Db	660	CTSLEYNIRVYCLHDTHDALKELVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPVPS	719
Qy	719	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWWQVEGDGQSFSI	778
Db	720	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWWQVEGDGQSFSI	779
Qy	779	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKISSLDPCCRRGADWRTLAQ	838
Db	780	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKISSLDPCCRRGADWRTLAQ	839
Qy	839	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	898
Db	840	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	899

RESULT 10

ADH71640

ID ADH71640 standard; protein; 899 AA.

XX

AC ADH71640;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21p SEQ ID NO:536.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.

PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR N-PSDB; ADH71639.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 536; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and

CC antilipaemic activity, and may have a use in gene therapy, and as a
 CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
 CC any of the 303 fully defined nucleotide sequences given in the
 CC specification. The polypeptide is useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease. The
 CC polypeptide, polynucleotide and antibody are useful in diagnosing,
 CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
 CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
 CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
 CC further used as hybridisation probes, in chromosome mapping, tissue
 CC typing, preventive medicine, and pharmacogenomics. The present sequence
 CC represents a NOVX polypeptide of the invention.

XX

SQ Sequence 899 AA;

Query Match 98.0%; Score 4695.5; DB 8; Length 899;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 887; Conservative 3; Mismatches 7; Indels 3; Gaps 3;

Qy	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Db	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Qy	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL	120
Db	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRRQVEKVFGL	120
Qy	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAE	180
Db	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAE	180
Qy	181	VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY	240
Db	181	VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNV-QKTACATLCPVDG	299
		: :	
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVHDRTVSSLLVSVDG	300
Qy	300	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	359
Db	301	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	360
Qy	360	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSIILTSQFQPVSIKPSKADNPHL	419
Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSIILTSQFQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLSTTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFEV	479
Db	421	LTIQPDLS-TTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFEV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTNFGGLRMLPNTGISLLIPDAIPRGKIYEIYTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTNFGGLRMLPNTGISLLIPDAIPRGKIYEIYTLH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGS	599

Db 540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGS 599
 Qy 600 WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVA 658
 || |||||
 Db 600 WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLRLLLFAPVA 659
 Qy 659 CTSLEYNIRVYCLHDT HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 718
 |||||
 Db 660 CTSLEYNIRVYCLHDT HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS 719
 Qy 719 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 778
 |||||
 Db 720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779
 Qy 779 NFNITKDTRFAELLA ESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRGGADWRTLAQ 838
 |||||
 Db 780 NFNITKDTRFAELLA ESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRGGADWRTLAQ 839
 Qy 839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C 898
 |||||
 Db 840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C 899

RESULT 11

ADH71632

ID ADH71632 standard; protein; 899 AA.

XX

AC ADH71632;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV211 SEQ ID NO:528.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
 KW obesity; diabetes; infectious disease; metabolic syndrome X;
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.

Qy	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL	120
Db	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGSEPTMEVRINVSRRQVEKVFGL	120
Qy	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE	180
Db	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE	180
Qy	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY	240
Db	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNV-QKTACATLCPVDG	299
		: :	
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVHDTVSSLLVSDG	300
Qy	300	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	359
Db	301	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	360
Qy	360	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSQFQPVSIKPSKADNPHL	419
Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSQFQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLSTTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Db	421	LTIQPDLSTTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGS	599
Db	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLLFAPVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLLFAPVA	659
Qy	659	CTSLEYNIRVYCLHDTDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDPVPS	718
Db	660	CTSLEYNIRVYCLHDTDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDPVPS	719
Qy	719	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSFSI	778
Db	720	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSFSI	779
Qy	779	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	838
Db	780	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	839
Qy	839	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAE	898
Db	840	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAE	899

RESULT 12

ADH71638

ID ADH71638 standard; protein; 899 AA.

XX

AC ADH71638;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21o SEQ ID NO:534.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.

PR 12-JUN-2002; 2002US-0387934P.

PR 12-JUN-2002; 2002US-0387960P.

PR 12-JUN-2002; 2002US-0388022P.

PR 12-JUN-2002; 2002US-0388096P.

PR 13-JUN-2002; 2002US-0389123P.

PR 14-JUN-2002; 2002US-0389118P.

PR 14-JUN-2002; 2002US-0389120P.

PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI MacLachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR N-PSDB; ADH71637.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 534; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC represents a NOVX polypeptide of the invention.

XX

SQ Sequence 899 AA;

Query Match 98.0%; Score 4694.5; DB 8; Length 899;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 887; Conservative 2; Mismatches 8; Indels 3; Gaps 3;

Qy	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Db	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Qy	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGL	120
Db	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGEPTMEVRINVSRQQVEKVFGL	120
Qy	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPP	180
Db	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPP	180
Qy	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASA	240
Db	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASA	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPPAPLNGGAFCEGQNV-QKTACATLCP	299
		: :	
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPPAPLNGGAFCEGQNVHDRTVSSLLV	300
Qy	300	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGP	359
Db	301	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGP	360
Qy	360	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSQFQPVSIKPSKADN	419
Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSQFQPVSIKPSKADN	420

Qy	420	LTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Db	421	LTIQPDLS-TTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSSWSRLKKQSCEGS	599
Db	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSSWSRLKKQSCEGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAATKRLKLLLFAPVA	659
Qy	659	CTSLEYNIRVYCLHDT HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS	718
Db	660	CTSLEYNIRVYCLHDT HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS	719
Qy	719	LWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI	778
Db	720	LWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI	779
Qy	779	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	838
Db	780	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	839
Qy	839	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C	898
Db	840	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C	899

RESULT 13

ADH71634

ID ADH71634 standard; protein; 899 AA.

XX

AC ADH71634;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21m SEQ ID NO:530.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.

PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI MacIachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR N-PSDB; ADH71633.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 530; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC represents a NOVX polypeptide of the invention.

XX

SQ Sequence 899 AA;

Query Match 98.0%; Score 4693.5; DB 8; Length 899;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 887; Conservative 2; Mismatches 8; Indels 3; Gaps 3;

Qy	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Db	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Qy	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGL	120
Db	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGSEPTMEVRINVSQQVEKVFGL	120
Qy	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE	180
Db	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE	180
Qy	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASA	240
Db	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASA	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNV-QKTACATLCPVDG	299
		: :	
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVHVRTVSSLLVSVDG	300
Qy	300	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDV	359
Db	301	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDV	360
Qy	360	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHL	419
Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLSTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Db	421	LTIQPDLS-TTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSSWSLRLKKQSCEGS	599
Db	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSSWSLHLKKQSCEGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA	659
Qy	659	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNRLSIHDPSS	718
Db	660	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNRLSIHDPSS	719
Qy	719	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSF	778

Db 720 LWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI 779

Qy 779 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 838
 |||

Db 780 NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839

Qy 839 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898
 |||

Db 840 KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 899

RESULT 14

ADH71648

ID ADH71648 standard; protein; 899 AA.

XX

AC ADH71648;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21t SEQ ID NO:544.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
 KW obesity; diabetes; infectious disease; metabolic syndrome X;
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNV-QKTACATLCPVDG	299
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVHVRTVSSLLVSDG	300
Qy	300	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	359
Db	301	SWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDV	360
Qy	360	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSGFQPVSIKPSKADNPHL	419
Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSGFQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLS-TTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSPTSEAEFEV	479
Db	421	LTIQPDLS-TTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSPTSEAEFEV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRIMIPNTGISLLIPPDAIPRGKIYEIYTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRIMIPNTGISLLIPPDAIPRGKIYEIYTLH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGS	599
Db	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAALKRLKLLFAPVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAALKRLKLLFAPVA	659
Qy	659	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPVSS	718
Db	660	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPVSS	719
Qy	719	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWWQVEGDGQSFSI	778
Db	720	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWWQVEGGGQSFSI	779
Qy	779	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPCCRRGADWRTLAQ	838
Db	780	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPCCRRGADWRTLAQ	839
Qy	839	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	898
Db	840	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	899

RESULT 15

ADH71630

ID ADH71630 standard; protein; 899 AA.

XX

AC ADH71630;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21k SEQ ID NO:526.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.

PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR N-PSDB; ADH71629.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 526; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a

Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA	AKRLKLLLFAPVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA	AKRLKLLLFAPVA	659
Qy	659	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIH	DVPSS	718
Db	660	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIH	DVPSS	719
Qy	719	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI		778
Db	720	LWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSI		779
Qy	779	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDP	PCRRGADWRTLAQ	838
Db	780	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDP	PCRRGADWRTLAQ	839
Qy	839	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	C	898
Db	840	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	C	899

Search completed: March 1, 2005, 08:56:42
Job time : 210.443 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2005, 08:42:47 ; Search time 53.8066 Seconds
(without alignments)
1245.848 Million cell updates/sec

Title: US-10-624-932-2
Perfect score: 4791
Sequence: 1 MAVRPGLWPALLGIVLAOWL.....AVAGLGQPDAGLFTVSEAE 898

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	4638	96.8	898	2	US-08-808-982-5	Sequence 5, Appli
2	4638	96.8	898	3	US-09-306-902A-5	Sequence 5, Appli
3	2815.5	58.8	557	2	US-08-808-982-6	Sequence 6, Appli
4	2815.5	58.8	557	3	US-09-306-902A-6	Sequence 6, Appli
5	2571.5	53.7	943	2	US-08-808-982-7	Sequence 7, Appli
6	2571.5	53.7	943	3	US-09-306-902A-7	Sequence 7, Appli
7	2337.5	48.8	769	4	US-09-949-016-10665	Sequence 10665, A
8	1895	39.6	886	4	US-09-969-532-16	Sequence 16, Appl
9	1886.5	39.4	897	4	US-09-969-532-14	Sequence 14, Appl
10	1878	39.2	900	4	US-09-969-532-12	Sequence 12, Appl
11	1869.5	39.0	911	4	US-09-969-532-10	Sequence 10, Appl

12	1259.5	26.3	655	4	US-09-969-532-32	Sequence 32, Appl
13	1251	26.1	666	4	US-09-969-532-30	Sequence 30, Appl
14	1242.5	25.9	669	4	US-09-969-532-28	Sequence 28, Appl
15	1234	25.8	680	4	US-09-969-532-26	Sequence 26, Appl
16	1124	23.5	552	4	US-09-969-532-8	Sequence 8, Appli
17	1115.5	23.3	563	4	US-09-969-532-6	Sequence 6, Appli
18	1107	23.1	566	4	US-09-969-532-4	Sequence 4, Appli
19	1098.5	22.9	577	4	US-09-969-532-2	Sequence 2, Appli
20	488.5	10.2	321	4	US-09-969-532-24	Sequence 24, Appl
21	480	10.0	332	4	US-09-969-532-22	Sequence 22, Appl
22	471.5	9.8	335	4	US-09-969-532-20	Sequence 20, Appl
23	463	9.7	346	4	US-09-969-532-18	Sequence 18, Appl
24	296.5	6.2	1172	1	US-08-313-288B-19	Sequence 19, Appl
25	296.5	6.2	1172	4	US-09-949-016-6333	Sequence 6333, Ap
26	294	6.1	102	2	US-08-808-982-8	Sequence 8, Appli
27	294	6.1	102	3	US-09-306-902A-8	Sequence 8, Appli
28	276	5.8	954	4	US-09-854-845-14	Sequence 14, Appl
29	276	5.8	1049	4	US-09-854-845-2	Sequence 2, Appli
30	276	5.8	1093	4	US-09-854-845-4	Sequence 4, Appli
31	276	5.8	1151	4	US-09-854-845-10	Sequence 10, Appl
32	272.5	5.7	939	4	US-09-854-845-16	Sequence 16, Appl
33	272.5	5.7	1034	4	US-09-854-845-6	Sequence 6, Appli
34	272.5	5.7	1078	4	US-09-854-845-8	Sequence 8, Appli
35	272.5	5.7	1136	4	US-09-854-845-12	Sequence 12, Appl
36	272.5	5.7	1170	4	US-09-657-472-2	Sequence 2, Appli
37	268.5	5.6	239	5	PCT-US93-01652-1	Sequence 1, Appli
38	268.5	5.6	1170	1	US-08-313-288B-20	Sequence 20, Appl
39	249.5	5.2	441	3	US-08-985-526-3	Sequence 3, Appli
40	249	5.2	479	4	US-09-270-767-46823	Sequence 46823, A
41	243	5.1	469	1	US-08-313-288B-15	Sequence 15, Appl
42	243	5.1	484	4	US-09-949-016-9698	Sequence 9698, Ap
43	242.5	5.1	1045	4	US-09-949-016-11112	Sequence 11112, A
44	238	5.0	218	3	US-08-985-526-1	Sequence 1, Appli
45	234.5	4.9	1395	3	US-09-540-245A-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1

US-08-808-982-5

; Sequence 5, Application US/08808982

; Patent No. 5939271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: Leonardo, E. David

; APPLICANT: Hink, Lindsay

; APPLICANT: Masu, Masayuki

; APPLICANT: Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

```

;      ZIP: 94104
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/808,982
;      FILING DATE:
;      CLASSIFICATION: 530
;      ATTORNEY/AGENT INFORMATION:
;      NAME: OSMAN, RICHARD A
;      REGISTRATION NUMBER: 36,627
;      REFERENCE/DOCKET NUMBER: UC96-217
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (415) 343-4341
;      TELEFAX: (415) 343-4342
;      INFORMATION FOR SEQ ID NO: 5:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 898 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: not relevant
;      TOPOLOGY: not relevant
;      MOLECULE TYPE: peptide
US-08-808-982-5

```

```

Query Match          96.8%; Score 4638; DB 2; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

```

```

Qy      1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
        |||
Db      1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGL 120
        |||
Db     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRQQVEKVFGL 120

Qy    121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAE 180
        |||
Db    121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAE 180

Qy    181 VEWLRNEDLVDP SLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY 240
        |||
Db    181 VEWLRNEDLVDP SLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAIVIVY 240

Qy    241 VNGGWSTWTEWSVCSASCGRGWQKRSRSC TNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300
        |||
Db    241 VNGGWSTWTEWSVCSASCGRGWQKRSRSC TNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300

Qy    301 WSPWSKWSACGLDCTHWR SRECSDPAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVA 360
        |||
Db    301 WSSWSKWSACGLDCTHWR SRECSDPAPRNGGEECRGADLDTRNCTSDLC LHTASCPEDVA 360

Qy    361 LYVGLIAVAVCLVLLLLVLILVYCRKKEGLDS DVADSSILTS GFQPVSIKPSKADNPHLL 420
        |||
Db    361 LYIGLVAVAVCLFLLLLALGLIYCRKKEGLDS DVADSSILTS GFQPVSIKPSKADNPHLL 420

```

Qy	421	TIQPDLSSTTTTTYQGS LCP RQDGPSPK FQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVS	480
Db	421	TIQPDLSSTTTTTYQGS LCSRQDGPSPK FQLSNGHLLSPLGSGRHTLHHSSPTSEAE DFVS	480
Qy	481	RLSTQNYFRSLPRGTSNM TYGTFN FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHK	540
Db	481	RLSTQNYFRSLPRGTSNM AYGTFN FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHK	540
Qy	541	PEDVRLPLAGCQTLLSP IVSCGPPGVLLTRPVILAMDHCGEPS PDSWSLRLKKQSCEGSW	600
Db	541	PEDVRLPLAGCQTLLSP VVSCGPPGVLLTRPVILAMDHCGEPS PDSWSLRLKKQSCEGSW	600
Qy	601	EDVLHLGEEAPSHLYYCQ LEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACT	660
Db	601	EDVLHLGEEPSHLYYCQ LEAGACYVFTEQLGRFALVGEALSVA ATKRLRLLLFAPVACT	660
Qy	661	SLEYNIRVYCLHDTHDAL KEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSI HDVPSSLW	720
Db	661	SLEYNIRVYCLHDTHDAL KEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSI HDVPSSLW	720
Qy	721	KSKLLVSYQEIPFYHIWNGT QRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINF	780
Db	721	KSKLLVSYQEIPFYHIWNGT QQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNINF	780
Qy	781	NITKDTRFAELLAL ESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL	840
Db	781	NITKDTRFAELLAL ESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSR GADWRTLAQKL	840
Qy	841	HLD SHLSFFASKPSPTAMI LNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC	898
Db	841	HLD SHLSFFASKPSPTAMI LNLWEARHFPNGNLGOLAAAVAGLGOPDAGLFTVSEAEC	898

US-09-306-902A-5

; GENERAL INFORMATION:

10 TITLE OF INVENTION: Netrin Receptors

CORRESPONDENCE ADDRESS:

; COMPUTER READABLE FORM:

Qy. 481 RLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHK 540
 |||||
 Db 481 RLSTQNYFRSLPRGTSNMAYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHK 540
 |||||
 Qy 541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCGWSW 600
 |||||:|||||
 Db 541 PEDVRLPLAGCQTLLSPVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCGWSW 600
 |||||:|||||
 Qy 601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAACKRLKLLLFAPVACT 660
 |||||:|||||
 Db 601 EDVLHLGEEAPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACT 660
 |||||:|||||
 Qy 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPVSSIW 720
 |||||
 Db 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPVSSIW 720
 |||||
 Qy 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINF 780
 |||||:|||||
 Db 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERINASTSDLACKLWVWQVEGDGQSFNINF 780
 |||||:|||||
 Qy 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL 840
 |||||:|||||
 Db 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL 840
 |||||:|||||
 Qy 841 HLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAEC 898
 |||||
 Db 841 HLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAEC 898
 |||||

RESULT 3

US-08-808-982-6

; Sequence 6, Application US/08808982

; Patent No. 5939271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: Leonardo, E. David

; APPLICANT: Hink, Lindsay

; APPLICANT: Masu, Masayuki

; APPLICANT: Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/808,982

; FILING DATE:

```
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-808-982-6
```

```
Query Match          58.8%; Score 2815.5; DB 2; Length 557;
Best Local Similarity 96.8%; Pred. No. 5.8e-259;
Matches 539; Conservative 2; Mismatches 15; Indels 1; Gaps 1;
```

```
Qy      343 NCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTS 402
        ||||| ||:|||||||||||||||||||||||||||||||||||||||||
Db       1 NCTSDLXVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTS 60

Qy      403 GFQPVSIKPSKADNPHELLTIQPDLSTTTTTYQGSILCPRQDGSPKPFQLTNGHLLSPILGGG 462
        |||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GFQPVSIKPSKADNPHELLTIQPDLSTTTTTYQGSILCPRQDGSPKPFQLTNGHLLSPILGGG 120

Qy      463 RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIP 522
        |||||||||||||||||||||||||||||||||||||||||||||||
Db      121 RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIP 180

Qy      523 PDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE 582
        |||||||||||||||||||||||||||||||||||||||||||||||
Db      181 PDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE 240

Qy      583 SPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS 642
        ||||| |||||||||||||||||||||||||||||||||||||||||||
Db      241 SPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS 300

Qy      643 VAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFK 702
        |||||||||||||||||||||||||||||||||||||||||||||||
Db      301 VAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHLX 360

Qy      703 DSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLAC 762
        ||||| || |||||||||||||||||||||||||||||||||||||||
Db      361 DSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLAC 420

Qy      763 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 822
        |||||||||||||||||||||||||||||||||||||||||||||||
Db      421 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 480

Qy      823 LDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG 882
        |||||||||||||||||||||||||||||||||||||||||||||||
Db      481 LDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG 540
```

Qy 883 LGQPDAGLFT-VSEAEC 898
| : |||||
Db 541 TXPAGRWLLSQCEAEC 557

RESULT 4

US-09-306-902A-6

; Sequence 6, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-306-902A-6

Query Match 58.8%; Score 2815.5; DB 3; Length 557;
Best Local Similarity 96.8%; Pred. No. 5.8e-259;
Matches 539; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

Qy 343 NCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 402
||||| |:|||||
Db 1 NCTSDLXVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 60


```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 943 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-808-982-7

```

```

Query Match          53.7%; Score 2571.5; DB 2; Length 943;
Best Local Similarity 53.3%; Pred. No. 2.5e-235;
Matches 504; Conservative 142; Mismatches 221; Indels 79; Gaps 16;

```

```

Qy      9 PALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAV 68
      |:| ||      ||| : : | | : ||||:|||| ||||| | |:|
Db     21 PSLAGI-----DSGAQ---GLPDSFPSAPAEQLPHFLLEPEDAYIVKNKPVLELHCRAF 70

Qy     69 PATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVA 128
      ||||:||||| | || : | | ::|| ||:| |||||:||||:|||||
Db     71 PATQIYFKCNGEWVSQKGHVTQESLDEATGLRIREVQIEVSRQQVEELFGLLEDYWCQCVA 130

Qy    129 WSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPPEGIPPAEVEWLNRNED 188
      |||||:||||| ||||:||||| | : ::| ||||:| ||||:|
Db    131 WSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPPEGVPVAEVEWLKRNED 190

Qy    189 LVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTW 248
      ::||: | | :| :|:|:| ||||:||||| ||||:| | |||||:|
Db    191 VIDPAQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRSTTATVIVYVNGGWSSW 250

Qy    249 TEWSVCSASCGRGWQKRSRCTNPAPLNGGAFCEGQNVQKTACATLCPVDGWSWSPWSKWS 308
      ||| || |||||:|:||||| |||| | :|||:|: ||||
Db    251 AEWSPCSNRCGRGWQKRTRCTNPAPLNGGAFCEGQACQKTACTTVCPVDGAWTEWSKWS 310

Qy    309 ACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCV---HSASGPE----- 357
      || :| ||||| | |:|| :| || ||:| || || : : |:
Db    311 ACSTECAHWSRECMAPPQNGGRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSRPLEPSG 370

Qy    358 DVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLSDVDADSS-ILTSGFQPVSIKPSKAD 415
      |||| || :|| | | :| :| || | :|: || || || ||: | ::
Db    371 DVALYAGLVVAVFVVLAVLMAVGIVYRRNCRDFDTDITDSSAALTGGFHPVNFKTARPS 430

```

Qy 416 NPHLL--TIQPDLSSTTTTTYQGS LCP RQDGPS PKFQLTNGHLLSPL----- 459
 || || : |||: : |:| : || : | :|| || ||
 Db 431 NPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDPLPSLKIKVYDSSSTIG 489
 Qy 460 -GGG-----RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGT 495
 | | | | | | | | : | :| : |||
 Db 490 SGAGLADGADLLGVLP PGTYPGDFSRDTHFLHLRS-----ASLGSQ-HLLGLPRDP 539
 Qy 496 SNMTYGT FNFLGGR LMI PNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLA-GCQTL 554
 |: ||| |||| || ||:||||: |||:|| |:|| :| | |||: | ||:
 Db 540 SSSVSGTFGCLGGRLTIPGTGVSL LVPNGAIPQ GKFDLYLRINKTEST-LPLSEGSQTV 598
 Qy 555 LSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGEEAPSHL 614
 ||| |:||| |:| |||:| : || | :|| |: :| |||:| || :
 Db 599 LSPSVTCGPTGLLLCRPV LTVPHCAEVIAGDWIFQLKTQAHQGHWEVVTLDEETLNTP 658
 Qy 615 YYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDT 674
 ||||| :|:: :||| : ||: | :| |||:| :|| ||||::||| ||
 Db 659 CYCQLEAKSCHILLDQLGT YVFTGESYSRS AVKRLQLAIFAPALCT SLEYSLRVYCLEDT 718
 Qy 675 HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFY 734
 |||||::||: || |:||: | |||||:|:| : |:||| |||||
 Db 719 PAALKEVLELERTLG GYLVEEPKTL LFKDSYHNLRLSLHDIPHAHWRSKLLAKYQEIPFY 778
 Qy 735 HIWNGTQRYLHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDT RFAELLAL 794
 |:|||:|: ||||| || :::: ||: | |||:| | : : :| ||
 Db 779 HVWNGSQKALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLDAL 837
 Qy 795 ESEAGVPAL--VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASK 852
 | | | :|| |||| |||| :|| | || ||| |||| :| :|::||:|
 Db 838 CSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLAQKLSMDRYLN YFATK 897
 Qy 853 PSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C 898
 ||| :||:|||| :|:|: ||:|: |: : : : : :
 Db 898 ASPTGVILDLEARQQDDGDLNSLASALEEMGKSEMLVAMTTD GDC 943

RESULT 6

US-09-306-902A-7

; Sequence 7, Application US/09306902A

; Patent No. 6277585

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 943 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-306-902A-7

```

```

Query Match          53.7%; Score 2571.5; DB 3; Length 943;
Best Local Similarity 53.3%; Pred. No. 2.5e-235;
Matches 504; Conservative 142; Mismatches 221; Indels 79; Gaps 16;

```

```

Qy      9 PALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAV 68
      |:| ||      ||| : : | | : ||||:|||| ||||| | |:|
Db     21 PSLAGI-----DSGAQ---GLPDSFPSAPAEQLPHFLLEPEDAYIVKNKPVLEHLCRAF 70

Qy     69 PATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGLLEEYWCQCVA 128
      ||||:|||||| | || : | | ::|| ||:| |||||::||||:|||||
Db     71 PATQIYFKCNGEWVSQKGHVTQESLDEATGLRIREVQIEVSRQQVEELFGLEDYWCQCVA 130

Qy    129 WSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNED 188
      |||||:|||||:||||| ||||:|||||| | : :| ||||:| ||||:|
Db    131 WSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPPEGVPVAEVEWLKNE 190

Qy    189 LVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTW 248
      ::||: | | :| :|:|:|:||||:|||||:|||||:| | | |||||:|
Db    191 VIDPAQDTNFLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRSTTATVIVYVNGGWSSW 250

Qy    249 TEWSVCSASCGRGWQKRSRSCITNPAPLNGGAFCEGQNVQKTACATLCPVDGWSWSPWSKWS 308
      ||| || |||||:|:||||| |||| |:||||:|: ||||
Db    251 AEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQACQKTACTTVCPVDGAWTEWSKWS 310

Qy    309 ACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCV---HSASGPE----- 357
      || :| ||||| | |:| | :| || |:| || | : : |
Db    311 ACSTECAHWSRECMAPPQNGGRDCSGTLLDSKNCTDGLCVLNQRTLNQKSRPLEPSG 370

Qy    358 DVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSS-ILTSGFQPVSIKPSKAD 415
      |||| | | :| | | :| | :| | | :| | | | | | :| :|
Db    371 DVALYAGLVVAVFVVLAVLMAVGIVYRRNCRDFDTDITDSSAALTGGFHPVNFKTARPS 430

```

Qy 416 NPHLL--TIQPDLSSTTTTQGSCLPRQDGPSPKFQLTNGHLLSPL----- 459
 || || : |||: : |:| : || : | :|| || ||
 Db 431 NPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDPLPSLKIKVYDSSTIG 489
 Qy 460 -GGG-----RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGT 495
 | | | | | : | : | : |||
 Db 490 SGAGLADGADLLGVLPPTYPGDFSRDTHFLHLRS-----ASLGSQ-HLLGLPRDP 539
 Qy 496 SNMTYGTNFNLGGRMLIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLA-GCQTL 554
 |: ||| |||| || ||:||||:| |||:| | |:| | :| | |||: | ||:
 Db 540 SSSVSGTFGCLGGRLTIPGTGVSLVPGAIPQGKFYDLYLRINKTEST-LPLSEGSQTV 598
 Qy 555 LSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGEEAPSHL 614
 ||| |:| | |:| |||:| : || | | :| | :| | |:| | || :
 Db 599 LSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIFQLKTQAHQGHWEVVTLDEETLNT 658
 Qy 615 YYCQLEASACYVFTEQLGRFALVGEALSVAARKLKLLEFAPVACTSLEYNIRVYCLHDT 674
 ||||| |:| : ||| : ||: | :| |||:| :|| | |||||:|||| | |
 Db 659 CYCQLEAKSCHILLDQLGTYVFTGESYSRSYSAVKRLQLAIFAPALCTSLEYSRLVYCLED 718
 Qy 675 HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSLIHDVPSSLWWSKLLVSYQEIPFY 734
 |||||:| |:| | |:| : | ||||| |||:| : | |:| | |||||
 Db 719 PAALKEVLELERTLGGYLVVEPKTLLFKDSYHNLRLSLHDIPHAHWRSLAKYQEIPFY 778
 Qy 735 HIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLAL 794
 |:| |:| |:| ||||| | :| :| ||:| | :| : : | | ||
 Db 779 HVWNGSQKALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLDAL 837
 Qy 795 ESEAGVPAL--VGPSAFKIPFLIRQKISSLDPCCRRGADWRTLAQKLHLDShLSFFASK 852
 | | | :| | ||| | ||| :|| | || ||| |||| :| :| :| :|
 Db 838 CSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLAQKLSMDRYLNYFATK 897
 Qy 853 PSPTAMILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAEC 898
 ||| :| :| ||| :| :| ||:| :| : : : : :|
 Db 898 ASPTGVILDLEWARQQDDGDLNSLASALEEMGKSEMLVAMTTDGDGDC 943

RESULT 7

US-09-949-016-10665

; Sequence 10665, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

```
; SEQ ID NO 10665
;   LENGTH: 769
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10665
```

Query Match 48.8%; Score 2337.5; DB 4; Length 769;
Best Local Similarity 56.1%; Pred. No. 3.4e-213;
Matches 432; Conservative 136; Mismatches 179; Indels 23; Gaps 7;

Qy	147	LRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSL	206
Db	3	LRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPVEDRNFYITIDHNL	62
Qy	207	VVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSSASCGRGWQKRS	266
Db	63	IIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRCGRGYQKRT	122
Qy	267	RSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DPA	326
Db	123	RTCTNPAPLNGGAFCEGQSVQKIACTTLC PVDGRWTPWSKWSTCGTECTHWRRECTAPA	182
Qy	327	PRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGL-I AVAVCLVLLLLVLILVYCR	385
Db	183	PKNGGKDCDGLVLQSKNCTDGLCMQTAPDSD DVALYVGIVIAVIVCLAISVVVALFVYRK	242
Qy	386	KKEGLDSDVADSSILTS G FQPVSIKPSKADNP HLLTIQPD LSTTTTTTYQGS LCP RQDGPS	445
Db	243	NHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPPDLTSAAAMYRGVPVYALHD-VS	298
Qy	446	PKFQLTN GHLLSPLGGGRHTLHHSS-----PTSEAEFEVSRLS---TQNYF-----R	489
Db	299	DKIPMTNSPILDPLPNLKIKVYNTSGAVTPQDDLSEFTSKLS PQMTQS LLENEALS LKNQ	358
Qy	490	SLPRGT--SNMTYGTENFLGGRLMIPNTGISLLIPPD AIPRGKIYEIYLT LHKPEDVRLP	547
Db	359	SLARQTDPSCTAFGSGFNSLGGHLIVPNSGVSLLI PAGAI PQGRVYEMYVTVHRKETMRPP	418
Qy	548	LAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE P SPDSWSLRLKKQSCEGSWEDVLHLG	607
Db	419	MDDSQTLLTPV VSCGPPGALLTRPVVLTMHHCADPN TEDWKI LLKNQA AQGWEDVVVVG	478
Qy	608	EEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLL FAPVACTSLEYNIR	667
Db	479	EENFTTPCYIQLD AEACHILTENLSTYALVGHSTTKAAAKRLKLAIFGPLCCSSLEYSIR	538
Qy	668	VYCLHDTHDALKEVVQLEKQLGGQLIQEPVLHFKDSYHNLRLSIHDVPSS LWSKLLVS	727
Db	539	VYCLDDTQDALKEILHLERQMGGQLLEEPKALHFKGSTHNLRLSIHDIAHSLWSKLLAK	598
Qy	728	YQEIPFYHIWNGTQRYLHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDTR	787
Db	599	YQEIPFYHVWSGSQRNLHCTFTLERFSLNTVELVCKLCVRQVEGEGQIFQLNCTVSE EPT	658
Qy	788	FAELLALESEAGVPALVGPSAFKIPFLIRQKI ISSLDPPCRGRGADWRTLAQKLHLD SHLS	847
Db	659	GIDLPLLD PANTITTTVTGPSAFSIP LPIROKLCSSLDAPOTRGHDWRMLAHKLNLD RYLN	718

Qy 848 FFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 897
 :||:| ||| :||:||||:|||| ||| : :| : : :| :
 Db 719 YFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMGRHETVVSLLAAEQ 768

RESULT 8

US-09-969-532-16

; Sequence 16, Application US/09969532
 ; Patent No. 6777232
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Scoville, John
 ; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides
 Encoding the Same
 ; FILE REFERENCE: LEX-0244-USA
 ; CURRENT APPLICATION NUMBER: US/09/969,532
 ; CURRENT FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: US 60/237,280
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 16
 ; LENGTH: 886
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-09-969-532-16

Query Match 39.6%; Score 1895; DB 4; Length 886;
 Best Local Similarity 41.2%; Pred. No. 6.1e-171;
 Matches 383; Conservative 157; Mismatches 277; Indels 112; Gaps 14;

Qy 8 WPALLGIVLAAWLRGS----GAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLL 63
 | || | | | : | : | | ||| : || : | : |
 Db 17 WLPWLGLCFWAAGTAAARGTDNGEALPESIPSA-PGTLPHFIEEPDDAYIIKSNPIAL 73
 Qy 64 VCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEW 123
 ||| || ||||| |||| | : || | : | |||| || ||| : |||| | : ||
 Db 74 RCKARPAMQIFFKCNGEWVHQNEHVSEETLDESSGLKRVREVFINVTRQQVEDFHGPEDYW 133
 Qy 124 CQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEW 183
 ||||| || : || : || : || ||||| : || : || || ||||| : |||||
 Db 134 CQCVAWSHLGTSKSRKASVRIAYLRKNFEQDPQGREVPPIEGMIVLHCRPPEGVPAAEVEW 193
 Qy 184 LRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNG 243
 | : || : | | : : : : ||||| : ||||| : ||||| : || : || : |
 Db 194 LKNEEPIDSEQDENIDTRADHNLIIRQARLSDSGNYTCMAANIVAKRRSLSATVVVYVDG 253
 Qy 244 GWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGSWSP 303
 | | : |||||
 Db 254 SWEVWSEWSVCSP----- 266
 Qy 304 WSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYV 363
 : | | | ||| : | |||| : || : : ||| || : | : |||
 Db 267 -----ECEHLRIRECTAPPPRNGGKFCEGLSQESENCTDGLCILGIENASDIALYS 317
 Qy 364 GLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHLL--T 421

```

Db      318 GL-GAAVVAVAVLVIGVTLYRRSQSDYGVDVIDSSALTGGFQTFNFKTVRQGNLLLLNSA 376
Qy      422 IQPDLSTTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLG----- 460
      :|||| | : || | :| || | | :| | :||
Db      377 MQPDL-TVSRTYSGPIC-LQD-PLDKELMTSESLFNPLSDIKVKVQSSFMVSLGVSERAE 433
Qy      461 -----GGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTFFNLGG 508
      | | : | : : : || | | | | | |
Db      434 YHGKNHSRTFPHGNNHSFSTMHPRNKM-PYIQNLS-----SLPTRTELRTTGVFHGLGG 486
Qy      509 RLMIPNTGISLLIPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLL 568
      ||::|||:|||| | | :|||::: | | | : ||| | :||| ::
Db      487 RLVPNTGVSLIPHGAIPENSWEIYMSINQGEPSLQSDGSEVLLSPEVTCGPPDMIV 545
Qy      569 TRPVILAMDHCGEPSPDWSLRLKKQSCGSEWEDVLHLGEEAPSHLYYCQLEASACYVFT 628
      | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      546 TTPFALTIPHCADVSEHWNHLKKRTQQGKWEEVMSVEDESTS--CYCLDLPFACHVLL 603
Qy      629 EQLGRFALVGEALSVAIAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQL 688
      : | : || | : : | :||: : | :| ||:|:| | : | | :|| | :
Db      604 DSFGTYALTGEPITDCAVKQLKVAVFGCMSCNSLDYNLRVYCVDNTPCAFQEVVSDERHQ 663
Qy      689 GGQLIQEPRVLHFKDSYHNRLSLIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTF 748
      |||::||:| | : :||:| :| ||: | : ||:| : | : : || | |
Db      664 GGQLLEPKLLHFKGNTFSLQISVLDIPFLWRIKPFTACQEVPPFSRVWCNSRQPLHCAF 723
Qy      749 TLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSA 808
      :||| :||: |:||: : ||: | : : | : | : : | | | |
Db      724 SLERYTPTTTQLSCKICIRQLKGHEQILQVQTSILESERETITFFAQEDSTFPAQTGPKA 783
Qy      809 FKIPFLIRQKIISLDPPCRAGDWRTLAQKLHLDLHLSFFASKPSPTAMILNLWEARHF 868
      |||: |||: | : | | :| ||: ||| : :||:|: : ||:| | | | |
Db      784 FKIPYSIRQRICATFDTNPAKGKDWQMLAQKNSINRNLSYFATQSSPSAVILNLWEARHQ 843
Qy      869 PNGNLSQLAAAVAGLGQPDAGLFTVSEAE 897
      :|:| || | : :| : | :||:
Db      844 HDGDLDSLACALEEIGRTHTKLSNISESQ 872

```

RESULT 9

US-09-969-532-14

; Sequence 14, Application US/09969532

; Patent No. 6777232

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides
Encoding the Same

; FILE REFERENCE: LEX-0244-USA

; CURRENT APPLICATION NUMBER: US/09/969,532

; CURRENT FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US 60/237,280

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 14

; LENGTH: 897
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-969-532-14

Query Match 39.4%; Score 1886.5; DB 4; Length 897;
Best Local Similarity 40.9%; Pred. No. 4e-170;
Matches 384; Conservative 157; Mismatches 276; Indels 123; Gaps 15;

```
Qy      8 WPALLGIVLAAWLGRS----GAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLL 63
      |  ||  |  |  |:  |  :  :|  |  |||:  ||:  ||:|:  |:  |
Db     17 WLPWLG--LCFWAAGTAAARGTDNGEALPESIPSA-PGTLPHFIEEPDDAYIIKSNPIAL 73

Qy     64 VCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGLLEEW 123
      |||  ||  |||||  |||||  |  :||  |  :  |  ||||  ||  |||:  ||||  |  |:||
Db     74 RCKARPAMQIFFKCNGEWVHQNEHVSEETLDESSGLKVREVFINVTRQQVEDFHGPEDYW 133

Qy    124 CQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEW 183
      |||||  ||:  ||:  ||  :||  |||||  |:  :||  :|  ||  |||||  |  |||||
Db    134 CQCVAWSHLGTSKSRKASVRIAYLRKNFEQDPQGREVP IEGMIVLHCRPPEGVPAAEVEW 193

Qy    184 LRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNG 243
      |:  ||:  :|  |  |:  :|:  ::  ||||:  |:  ||||:  |  ||||:  ||  |  |:  ||:  |
Db    194 LKNEPIDSEQDENIDTRADHNLIRQARLSDSGNYTCMAANIVAKRRSLSATVVVYVDG 253

Qy    244 GWSTWTEWSVCSASCGRGWQKRSRCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSP 303
      |  |:  |||||
Db    254 SWEVWSEWSVCSP----- 266

Qy    304 WSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCV-----HS 352
      :|  |  |  |||:  |  ||||:  |:  |  ::  ||  ||:  |
Db    267 -----ECEHLRIRECTAPPPRNGGKFCEGLSQESENCTDGLCILDKKPLHEIKPQS 317

Qy    353 ASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSIILTSGFQPVSIKPS 412
      |:  |||  ||  ||  |  :|:  :  :|  :  ||  |||  ||  |||  :  |
Db    318 IENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSQSDYGVDVIDSSALTGGFQTFNFKTV 376

Qy    413 KADNPHELL--TIQPDLSSTTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLG----- 460
      :  |  ||  :|||  |  :  ||  |  :|  ||  |  :|  |  :||
Db    377 RQGNLILLNSAMQPDL-TVSRTYSGPIC-LQD-PLDKELMTESLFLNPLSDIKVKVQSSF 433

Qy    461 -----GGRHTLHHSSPTSEAEFFVSRLSTQNYFRSLPRGTSN 497
      |  |:  |  |  ::  ::  ||  |||  |
Db    434 MVSLGVSERA EYHGKNHSRTFPHGNNHSFSTMHPRNKM-PYIQNLS-----SLPTRTEL 486

Qy    498 MTYGTNFNFGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSP 557
      |  |  |  |||||:  ||||:  ||||  |||  :|||:  ::  |  |  |  :  ||||
Db    487 RTTGVFHGLGGRLVMPNTGVSLIPHGAIP EENSWEIYMSINQGEPSLQSDGSEVLLSP 545

Qy    558 IVSCGPPGVLLTRPVILAMDHCGEPSPDSSWLSRLKKQSCEGSWEDVLHLGEEAPSHLYYC 617
      |:  ||||  ::  |  |  :  ||  |  :  ||:  ||:  :|  ||:  |  :  :|  |  ||
Db    546 EVTCGPPDMIVTTPFALTIPHCADVSS EHWNIHLKKRTQQGKWEVMSVEDESTS--CYC 603

Qy    618 QLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDA 677
      |:  ||:  |  :  |  :||  ||  :  |  |:  |:  :|  ::  |  ||:  |:  ||||:  :|  |
Db    604 LLDPFACHVLLDSFGTYALTGEPITDCAVKQLKVAVFGCMSCNSLDYNLRVYCDVNTPCA 663
```


Qy 678 LKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKSLLVSYQEIPFYHIW 737
 :||| |: |||::||::||| : :||:: |:| ||: | : ||:| :|
 Db 664 FQEVVSDERHQGGQLLEPKLLHFKGNTFSLQISVLDIPPFLWRIKPFTACQEVPFPSRVW 723
 Qy 738 NGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESE 797
 :: ||| |:||| :||: |:||: : |:| | : :| : | : :
 Db 724 CSNRQPLHCAFSLERYTPTTTQLSCKICIRQLKGHEQILQVQTSILESERETITFFAQED 783
 Qy 798 AGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTA 857
 : || || ||||: |||:| :: | | :| ||: |||| :: :||:|:: ||:|
 Db 784 STFPAQTGPKAFKIPYSIRQRICATFDTPNAGKGDWQMLAQKNSINRNLSYFATQSSPSA 843
 Qy 858 MILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAE 897
 :||||||| :|:| || |: :|: | :|::
 Db 844 VILNLWEARHQHDGDLDSLACALEEIGRTHTKLSNISESQ 883

RESULT 10

US-09-969-532-12

; Sequence 12, Application US/09969532

; Patent No. 6777232

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides
 Encoding the Same

; FILE REFERENCE: LEX-0244-USA

; CURRENT APPLICATION NUMBER: US/09/969,532

; CURRENT FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US 60/237,280

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 900

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-969-532-12

Query Match 39.2%; Score 1878; DB 4; Length 900;

Best Local Similarity 40.6%; Pred. No. 2.6e-169;

Matches 383; Conservative 157; Mismatches 277; Indels 126; Gaps 15;

Qy 8 WPALLGIVLAAWLGRS----GAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLL 63
 | || | | |: | : :| | | |||: ||:| ||:|: |:|
 Db 17 WLPWLG--LCFWAAGTAAARGTDNGEALPESIPSA-PGTLPHFIEEPDDAYIIKSNPIAL 73
 Qy 64 VCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEW 123
 ||| || ||||| ||| |:| |:| | ||| || |||: ||| | |:|
 Db 74 RCKARPAMQIFFKCNGEWVHQNEHVSEETLDESSGLKVREVFINVTRQQVEDFHGPEDYW 133
 Qy 124 CQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEW 183
 ||||| ||:|:|:| :|| |||||:| :|| :| || |||||:| |||||
 Db 134 CQCVAWSHLGTSKSRKASVRIAYLRKNFEQDPQGREVPPIEGMIVLHCRPPEGVPAAEVEW 193
 Qy 184 LRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNG 243
 |:|:| :| | |: :|:|:| |||:| |||:| |||:| || |:|:|:


```

; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides
Encoding the Same
; FILE REFERENCE: LEX-0244-USA
; CURRENT APPLICATION NUMBER: US/09/969,532
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 60/237,280
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 911
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-969-532-10

```

```

Query Match          39.0%; Score 1869.5; DB 4; Length 911;
Best Local Similarity 40.3%; Pred. No. 1.7e-168;
Matches 384; Conservative 157; Mismatches 276; Indels 137; Gaps 16;

```

```

Qy      8 WPALLGIVLAAWLGRS----GAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLL 63
      |  || | | | :  |  :  : | | | |||: ||:| ||:| : |
Db     17 WLPWLG--LCFWAAGTAAARGTDNGEALPESIPSA-PGTLPHFIEEPDDAYIIKSNPIAL 73

Qy     64 VCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEW 123
      ||| || ||||| ||||| | :|| | : | ||||  || |||: ||||  | |:||
Db     74 RCKARPAMQIFFKCNGEWVHQNEHVSEETLDESSGLKVREVFINVTRQQVEDFHGPEDYW 133

Qy    124 CQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEW 183
      ||||| ||: ||: || :|| | |||||: | :|| :| || | |||||: | ||||
Db    134 CQCVAWSHLGTSKSRKASVRIAYLRKNFEQDPQGREVPPIEGMIVLHCRPPEGVPAAEVEW 193

Qy    184 LRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNG 243
      |:|: :| | :| :|: |||||: |||||: |||||: || | |:|: |
Db    194 LKNEEPIDSEQDENIDTRADHNLIIIRQARLSDSGNYTCMAANIVAKRRSLSATVVVYVDG 253

Qy    244 GWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGSWSP 303
      | | : |||||
Db    254 SWEVWSEWSVCSP----- 266

Qy    304 WSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCV-----HS 352
      :| | | ||: | ||||: |:| : : || | ||: |
Db    267 -----ECEHLRIRECTAPPPRNGGKFCEGLSQESENCTDGLCILDKKPLHEIKPQS 317

Qy    353 ASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTSGFQPVSIKPS 412
      |:||| ||  || | :|: : :| | :  || ||| || ||| : |
Db    318 IENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSQSDYGVDDVIDSSALTGGFQTFNFKTV 376

Qy    413 KA-----DNPHLL--TIQPDLSSTTTTYQGSCLPRQDGPSPKFQLTNGHLL 456
      :  |  || : |||| | : || | :| || | | :| |
Db    377 RQAKNIMELMIQEKSFGNSLLNSAMQPD-LTVSRTYSGPIC-LQD-PLDKELMTESSLF 433

Qy    457 SPLG-----GGRHTLHHSSPTSEAEFVSRLS 483
      :||  | | :  | : : : ||
Db    434 NPLSDIKVKVQSSFMVSLGVSEAEYHGKNHSRTFPHGNNHSFSTMHPRNKM-PYIQNLS 492

Qy    484 TQNYFRSLPRGTSNMITYGTNFGRLMIPNTGISLLIPDAIPRGKIYEIYLTTLHKPED 543

```


Qy 282 EGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDLDT 341
 Db 36 -----ECEHLRIRECTAPPPRNGGKFCEGLSQES 64

Qy 342 RNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILT 401
 Db 65 ENCTDGLCILGIENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSQSDYGV DVIDSSALT 123

Qy 402 SGFQPVSIKPSKADNPHLL--TIQPD LSTTTTTYQGS LCP RQDGPSPKFQLTNGHLLSPL 459
 Db 124 GGFQTFNFKTVRQGN SLLLNSAMQPD L-TVSR TYSGPIC-LQD-PLDKELMT ESSLFNPL 180

Qy 460 G-----GGRHTLHHSSPTSEAE EFVSRLSTQN 486
 Db 181 SDIKVKVQSSFMVSLGV SERAEYHGKNHSRTFPHGNNHSFSTMHP RNKM-PYIQNLS--- 236

Qy 487 YFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRL 546
 Db 237 ---SLPTRTELRTTG VFGHLGGRLVMPNTGV SLLIPHGAIPEENSWEIYMSINQGE P-SL 292

Qy 547 PLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSD SWSLRLKKQSCGE SWEDVLHL 606
 Db 293 QSDGSEVLLSPEVTCGPPDMIVTTPFALTIPHCADVSSEHWN IHLKKRTQQGKWE EVMSV 352

Qy 607 GEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLL FAPVACTSLEYNI 666
 Db 353 EDESTS--CYCLDPFACHVLLDSFGTYALTGEPI TDCAVKQLK VAVFGCMSCNSLDYNL 410

Qy 667 RVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKS KLLV 726
 Db 411 RVYCVDNTPCAFQEVVS DERHQQGQLLEEPKLLHFKGNTFSLQISVLDIPPFLWR IKPFT 470

Qy 727 SYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDT 786
 Db 471 ACQEVPPFSRVWC SNRQPLHCAFSLERYTP TTTQLSCKICIRQLKGHEQILQVQTSILESE 530

Qy 787 RFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTL AQKLHLD SHL 846
 Db 531 RETITFFAQEDSTFPAQTGPKAFKIPYSIRQRICATFDT PNAKGKDWQMLAQKNSINRNL 590

Qy 847 SFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 897
 Db 591 SYFATQSSPSAVILNLWEARHQHDGDLDSLACALEEIGRTHTKLSNISESQ 641

RESULT 13

US-09-969-532-30

; Sequence 30, Application US/09969532

; Patent No. 6777232

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; TITLE OF INVENTION: No. 6777232e1 Human Membrane Proteins and Polynucleotides
 Encoding the Same

; FILE REFERENCE: LEX-0244-USA

; CURRENT APPLICATION NUMBER: US/09/969,532

```
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 60/237,280
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 666
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-969-532-30
```

```
Query Match          26.1%; Score 1251; DB 4; Length 666;
Best Local Similarity 36.4%; Pred. No. 8.5e-110;
Matches 263; Conservative 128; Mismatches 215; Indels 116; Gaps 12;
```

```
Qy      222 VAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPLNGGAFC 281
      :| ||||:|||| || |:|||:| | |:|||||
Db      1  MAANIVAKRRSLSATVVVYVDGSWEVWSEWSVCSP----- 35

Qy      282 EGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDT 341
      :| | | |||: | ||||: |:| :
Db      36 -----ECEHLRIRECTAPPPRNGGKFCEGLSQES 64

Qy      342 RNCTSDLCV-----HSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGL 390
      ||| ||: | |:||| || | | |:|: : :| | :
Db      65 ENCTDGLCILDKKPLHEIKPQSIENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSQSDY 123

Qy      391 DSDVADSSILTSGFQPVSIKPSKADNPHLL--TIQPDLSSTTTTTYQGSCLPRQDGPSPKF 448
      || ||| || ||| : | : | || :||| | : || | :| || | |
Db      124 GVDVIDSSALTGGFQTFNFKTVRQGNLSLLNSAMQPDL-TVSRTYSGPIC-LQD-PLDKE 180

Qy      449 QLTNGHLLSPLG-----GGRHTLHHSSPTSEA 475
      :| | :|| | | : | :
Db      181 LMTESLFFNPLSDIKVKVQSSFMVSLGVSERAEYHGKNHSRTFPHGNNHSFSTMHPRNKM 240

Qy      476 EEFVSRLSTQNYFRSLPRGTSNMITYGTFFNLGGRLLMI PNTGISLLIPDAIPRGKIYEIY 535
      :: || || | | | | ||||:||||:|||| || | :|||
Db      241 -PYIQNLS-----SLPTRTELRTTGVFGHLGGRLVMPNTGVSLLI PHGAIPEENSWEIY 293

Qy      536 LTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDPSWSLRLKKQS 595
      ::::: | | | : |||| |:|||| ::| | | : || : | : |: : |||:
Db      294 MSINQGEP-SLQSDGSEVLLSPEVTCGPPDMI VTTTPFALTIPHCADVSSEHWNHLLKKRT 352

Qy      596 CEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLF 655
      :| |||: : : :| | | : ||| : | :|| || : | |:| : :|
Db      353 QQGKWEVMSVEDESTS--CYCLLPFACHVLLDSFGTYALTGEPITDCAVKQLKVAVFG 410

Qy      656 PVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDV 715
      ::| ||:|:|:|:|: :| | :||| |: ||||:|:|:|:|: : :|:|: | :
Db      411 CMSCNSLDYNLRVYCV DNTPCAFQEVVSDERHQGGQLLEPKLLHFKGNTFSLQISVLDI 470

Qy      716 PSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQS 775
      | ||: | : ||:| :| : || | |:||| :|:| :|:|: |:| |
Db      471 PPFLWRIKPFTACQEVFFSRVWC SNRQPLHCAFS LERYTPTTTQLSCKICIRQLKGHEQI 530

Qy      776 FSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPCCRRGADWRT 835
      : :| : | : : || || ||||: |||:| : : | | :| ||:
```

```

Db      531 LQVQTSILESERETITFFAQEDSTFPAQTGPKAFKIPYSIRQRICATFDTPNAKGKDWQM 590
Qy      836 LAQKLHLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSE 895
      |||| :: :||:||||: ||:|:||||||| :|:| || |: :|: | :||
Db      591 LAQKNSINRNLSYFATQSSPSAVILNLWEARHQHDGDLDSLACALEEIGRTHTKLSNISE 650
Qy      896 AE 897
      ::
Db      651 SQ 652

```

RESULT 14

US-09-969-532-28

```

; Sequence 28, Application US/09969532
; Patent No. 6777232
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides
Encoding the Same
; FILE REFERENCE: LEX-0244-USA
; CURRENT APPLICATION NUMBER: US/09/969,532
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 60/237,280
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 669
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-969-532-28

```

```

Query Match          25.9%; Score 1242.5; DB 4; Length 669;
Best Local Similarity 36.1%; Pred. No. 5.5e-109;
Matches 262; Conservative 128; Mismatches 216; Indels 119; Gaps 12;

```

```

Qy      222 VAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCASCGRGWQKRSRSCCTNPAPLNGGAFC 281
      :| ||||:|||| || |:||||:| | |:|||||
Db      1 MAANIVAKRRSLSATVVVYVDGSWEVWSEWSVCSP----- 35

Qy      282 EGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDT 341
      :| | | |||: | ||||: |:| ::
Db      36 -----ECEHLRIRECTAPPPRNGGKFCEGLSQES 64

Qy      342 RNCTSDLCVHSASGPEDVALYVGLIHAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILT 401
      ||| ||: |:||| || || |:|: : :| | : || ||| ||
Db      65 ENCTDGLCILGIENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSQSDYGVDVIDSSALT 123

Qy      402 SGFQPVSIKPSKA-----DNPHLL--TIQPDLTSTTTTYQGSICPRQDGPS 445
      ||| :| : | || :|||| |:| |||
Db      124 GGFQTFNFKTVRQAKNIMELMIQEKSFGNLSLLNSAMQPDL-TVSRTYSGPIC-LQD-PL 180

Qy      446 PKFQLTNGHLLSPLG-----GGRHTLHHSSPT 472
      | :| | :|| | | :|
Db      181 DKELMTESSLFNPLSDIKVKVQSSFMVSLGVSEAEYHGKNHSRTFPHGNNHSFSTMHPR 240

```

Qy 473 SEAEFVSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIPPDAPRGKIY 532
 :: :: || || | | | | ||||::|||:|||| ||| :
 Db 241 NKM-PYIQNLS-----SLPTRTELRTTGVFGHLGGRLVMPNTGVSLLIPHGAIPPENSW 293

Qy 533 EIYLTLLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLK 592
 |||:::| | | : ||| |:||| ::| | : || : | : |:: ||
 Db 294 EIYMSINQGEP-SLQSDGSEVLLSPEVTCGPPDMIVTTPFALTIPHCADVSEHWNHLK 352

Qy 593 KQSCGSEWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAANKRLKLL 652
 |: :| |:|: : :|: | || |: ||:| : | :|| || :: | |:|:
 Db 353 KRTQQGWEEVMSVEDESTS--CYCLDLPFACHVLLDSFGTYALTGEPITDCAVKQLKVA 410

Qy 653 LFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSI 712
 :| ::| |:|:||||: :| | :|| |: |||::||:|||| : :|::|:
 Db 411 VFGCMSCNSLDYNLRVYCVDNTPCAFQEVVSDERHQGGQLLEPKLLHFKGNTFSLQISV 470

Qy 713 HDVPSSLWWSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGD 772
 |:| ||: | : ||:| | : : || | |:| | |:|: |:|: : |:|
 Db 471 LDIPPFLWRIKPFTACQEVFSPVWCSNRQPLHCAFSLERYTPTTQLSCKICIRQLKGH 530

Qy 773 GQSFSINFNITKDTFAELLALESEAGVPALVGPSAFKIPFLIRQKISSLDPCCRAGD 832
 | : :| : | : : || || ||||: |||:| :: | | :| |
 Db 531 EQILQVQTSILESERETITFFAQEDSTFPAQTGPKAFKIPYSIRQRICATFDTPNAKGKD 590

Qy 833 WRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFT 892
 |: |||| :: :|:|:|: ||:|:||||||| :|:| || |: :|: |
 Db 591 WQMLAQKNSINRNLSYFATQSSPSAVILNLWEARHQHDGDLDSLACALEEIGRTHTKLSN 650

Qy 893 VSEAE 897
 :||:
 Db 651 ISESQ 655

RESULT 15

US-09-969-532-26

; Sequence 26, Application US/09969532

; Patent No. 6777232

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides
 Encoding the Same

; FILE REFERENCE: LEX-0244-USA

; CURRENT APPLICATION NUMBER: US/09/969,532

; CURRENT FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US 60/237,280

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 26

; LENGTH: 680

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-969-532-26

Query Match 25.8%; Score 1234; DB 4; Length 680;

Best Local Similarity 35.7%; Pred. No. 3.6e-108;

Matches 263; Conservative 128; Mismatches 215; Indels 130; Gaps 13;

```
Qy      222 VAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFC 281
      :| ||||:| | | |:| | | | | |:| | | | |
Db      1  MAANIVAKRRSLSATVVVYVDGSWEVWSEWSVCSP----- 35

Qy      282 EGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDT 341
      :| | | | | | | | | | | | | | | | | | | | | | |
Db      36 -----ECEHLRIRECTAPPPRNGGKFCEGLSQES 64

Qy      342 RNCTSDLCV-----HSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGL 390
      ||| ||: | | | | | | | | | | | | | | | | | |
Db      65 ENCTDGLCILDKKPLHEIKPQSIENASDIALYSLG-GAAVVAVAVLVIGVTLYRRSQSDY 123

Qy      391 DSDVADSSILTSGFQPVSIKPSKA-----DNPHLL--TIQPDLTSTTTTTYQ 434
      || ||| || ||| : | : | | | | | | | | | | | |
Db      124 GVDVIDSSALTGGFQTFNFKTVRQAKNIMELMIQESFGNSLLNSAMQPD-LTVSRTYS 182

Qy      435 GSLCPRQDGSPKFLTNHLLSPLG-----G 461
      | :| || | | | | | | | | | | | | | | | |
Db      183 GPIC-LQD-PLDKELMTESSLFNPLSDIKVKVQSSFMVSLGVSEAEYHGKNHSRTFPHG 240

Qy      462 GRHTLHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLI 521
      |: | : : : || | | | | | | | | | | | | | | |
Db      241 NNHSFSTMHPRNKM-PYIQNLS-----SLPTRTELRTGVFGHLGGRLVMPNTGVSLLI 293

Qy      522 PPDAIPRGKIYEIYLTTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE 581
      | ||| :| | | | | | | | | | | | | | | | | |
Db      294 PHGAIPEENSWEIYMSINQGEPLQSDGSEVLLSPEVTCGPPDMIVTTPFALTIPHCAD 352

Qy      582 PSPDSWSRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 641
      | : | : : | | | : : | | | | | | | | | | | |
Db      353 VSSEHWNHLKKRTQQGKWEVMSVEDESTS--CYCLDPFACHVLDSFGTYALTGEPI 410

Qy      642 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF 701
      : | | | : | | : | | | | | | : | | | | | | | |
Db      411 TDCAVKQLKVAVFGCMSCNSLDYNLRVYCDNTPCAFQEVVSDERHQGGQLLEPKLLHF 470

Qy      702 KDSYHNLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 761
      | : : | : | | | | | : | | | | | : | | | | | |
Db      471 KGNTFSLQISVLDIPPFLWRIKPFTACQEVFSPRVWCSNRQPLHCAFSLERYTPTTTQLS 530

Qy      762 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 821
      || : : | : | | : : | | : : | | | | | | | |
Db      531 CKICIRQLKGHEQILQVQTSILESERETITFFAQEDSTFPAQTGPKAFKIPYSIRQRICA 590

Qy      822 SLDPPCRRGADWRTLQKLHLDHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 881
      : | | : | | | | | | : : | | | | | | | | | |
Db      591 TFDTPNAKGKDQWMLAQKNSINRNLSYFATQSSPSAVILNLWEARHQHDGDLDSLACALE 650

Qy      882 GLGQPDAGLFTVSEAE 897
      :| : | : | : |
Db      651 EIGRTHTKLSNISESQ 666
```

Search completed: March 1, 2005, 09:05:46
Job time : 57.8066 secs

OM protein - protein search, using sw model

Run on: March 1, 2005, 08:46:18 ; Search time 35.0558 Seconds
 (without alignments)
 2464.715 Million cell updates/sec

Title: US-10-624-932-2
 Perfect score: 4791
 Sequence: 1 MAVRPGLWPALLGIVLAAWL.....AVAGLGQPDAGLFTVSEAEC 898

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : PIR_79:*
 1: pir1:*
 2: pir2:*
 3: pir3:*
 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8						
Result	Query					
No.	Score	Match	Length	DB	ID	Description
1	977	20.4	919	2	T32541	unc-5 protein - Ca
2	977	20.4	947	1	B44294	unc-5 protein, lon
3	298.5	6.2	1584	2	T00026	brain-specific ang
4	296.5	6.2	1172	1	TSHUP2	thrombospondin 2 p
5	293	6.1	1074	2	JC5928	semaphorin F precu
6	293	6.1	1172	2	A42587	thrombospondin 2 p
7	276	5.8	1444	2	T18856	angiogenesis inhib
8	275	5.7	984	2	T00326	hypothetical prote
9	275	5.7	1522	2	T00028	brain-specific ang
10	274.5	5.7	1572	2	T00027	brain-specific ang
11	270.5	5.6	1170	2	A40558	thrombospondin 1 p
12	268.5	5.6	1170	1	TSHUP1	thrombospondin 1 p
13	263	5.5	1178	1	A39804	thrombospondin pre

14	243	5.1	469	1	S29126	properdin precurs
15	229	4.8	437	2	S05478	properdin - mouse
16	226	4.7	254	2	T15952	hypothetical prote
17	221.5	4.6	1265	1	A37967	neural cell adhesi
18	215	4.5	788	2	T25061	hypothetical prote
19	215	4.5	1651	2	T14160	transmembrane rece
20	208	4.3	1612	2	T30805	dutt1 protein - mo
21	191.5	4.0	1344	2	T14316	rig-1 protein - mo
22	191	4.0	1863	2	S46217	protein-tyrosine-p
23	188.5	3.9	957	2	T15976	hypothetical prote
24	186	3.9	423	2	T29549	hypothetical prote
25	181.5	3.8	1273	2	T42405	sax-3 protein - Ca
26	181	3.8	1736	2	A47747	tight junction pro
27	178	3.7	1745	2	A46431	tight junction-ass
28	175	3.7	1907	2	S50893	protein-tyrosine-p
29	172	3.6	837	2	T00355	hypothetical prote
30	169.5	3.5	934	1	A34372	complement C6 prec
31	168.5	3.5	152	2	D89753	protein F11C7.2 [i
32	168	3.5	860	2	T16892	hypothetical prote
33	162.5	3.4	654	2	T29247	hypothetical prote
34	161	3.4	1501	2	I58148	protein-tyrosine-p
35	159.5	3.3	951	2	T00017	gene ADAMTS-1 prot
36	159	3.3	805	2	T34212	hypothetical prote
37	158.5	3.3	2165	2	T21371	hypothetical prote
38	157	3.3	550	2	T47158	hypothetical prote
39	156.5	3.3	807	2	A38152	F-spondin - rat
40	156	3.3	584	1	C8HUA	complement C8 alph
41	155	3.2	1499	2	I50212	protein-tyrosine-p
42	155	3.2	1898	2	S46216	leukocyte antigen-
43	150.5	3.1	1437	2	T31093	probable protein-t
44	149.5	3.1	712	2	A45638	immunodominant mic
45	148.5	3.1	206	2	A45517	coccidiosis-relate

ALIGNMENTS

RESULT 1

T32541

unc-5 protein - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T32541

R;Latreille, P.

submitted to the EMBL Data Library, December 1997

A;Description: The sequence of C. elegans cosmid B0273.

A;Reference number: Z21187

A;Accession: T32541

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-919 <LAT>

A;Cross-references: UNIPROT:O44171; EMBL:AF036698; PIDN:AAB88355.1;

GSPDB:GN00022; CESP:B0273.4a

A;Experimental source: strain Bristol N2; clone B0273

C;Genetics:

A;Gene: unc-5; CESP:B0273.4a

A;Map position: 4

A;Introns: 41/3; 108/1; 142/3; 201/1; 323/2; 553/1; 858/3
 C;Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology;
 thrombospondin type 1 repeat homology

Query Match 20.4%; Score 977; DB 2; Length 919;
 Best Local Similarity 28.7%; Pred. No. 2.3e-62;
 Matches 265; Conservative 168; Mismatches 379; Indels 110; Gaps 31;

Qy	49	EPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIER--STDGSSGLPTMEVRI	106
		: : :: : : : : : : : : :: :	
Db	9	QPKSGYVIRNKPLRLQCRANHATKIRYKCSSKWID--DSRIEKLIGTDSTSGVGIDASV	66
Qy	107	NVSRQQVEKVFGLEEYWCQCVAWSSSG-----TTKSQKAYIRIARLRKNFEQEPLAKEVS	161
		:: : : : : : : : : : :	
Db	67	DISRIDVDTSGHVDAFQCQCYA---SGDDDQDVVASDVATVHLAYMRKHFLKSPVAQRVQ	123
Qy	162	LEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTC	221
		: : : : : : :	
Db	124	EGTTLQLPCQAPESDPKAELTWYKDGVVVQP--DANVIRASDGSIMSAAARLSDSGNYTC	181
Qy	222	VAKNIVARRRSASAAVIVYVNGGWSTWTEW-SVCSASCG-----RGWQKR	265
		: : : : :	
Db	182	EATNVANSRKTDPEVQIYVDGGWSESWPWIGTCHVDCPLLRRQHAHRIRDPHDVLPHQRR	241
Qy	266	SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWKSWSACGLDCTHWSRECSDP	325
		: : : : : : : :	
Db	242	TRTCNNPAPLNDGEYCKGEEEMTRSKVPCKLDGGWSSWSWSACSSSCHRYRTRACTVP	301
Qy	326	APRNGGEECQGTDLDRNCTSDLCVHSASG--PEDVALYVGLIAVAVCLVLLLLVLILVY	383
		: : : : : : : : : : :	
Db	302	PPMNGGQPCFGDDLMTQECPAQLCTADSSRIVISDTAVYGSVASIFIVASFILAILAMFC	361
Qy	384	CR-----KKEGLSDSDVADSSILTSQFQVPSIKPSKADNPHELLTI-----	422
		: : : : : : : : : : : : : : :	
Db	362	CKRGNKSKSKPLKPQKMNSEKAGGIYYS---EPPGVRRLLEHQHGTLLGEKISSCSQYF	418
Qy	423	-QPDLSSTTTT-----TYQGSGLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPT-SE	474
		: :	
Db	419	EPPPLPHSTTLRSGKSAFSGYSSTRNAGSRAALIQECSSSSSSGSGGKRTMLRTSSSNCSD	478
Qy	475	AEEFVSRLSTQNYFRSLPRGTS-NMTYGTGFNGLGGRMLIPNTGISLLIPPDAPRGKIYE	533
		: : : : : : : : : :	
Db	479	DDNYATLYDYMEDKSVLGLDTSQNIVAAQIDSNGARLSLSKSGARLIVPELAVEGEKM--	536
Qy	534	IYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGV-----LLTRPVILAMDHCGEPSP-D	585
		: : : : : : : : :	
Db	537	LYLAVSDTLTDQPHLKPIESALSPVIVIGQCDVSMASHDNILRRPVVVSFRHCASTFPRD	596
Qy	586	SWSLRLKKQSCEGS-WEDVLHLGEEAPSHLYYCQLEASA-----CYVFTEQLGRFAL	636
		: : : : : : :	
Db	597	NWQFTL--YADEGSGWQKAVTIGEENLNTNMFVQFEQPGKKNDGFGWCHVMTYSLARLML	654
Qy	637	VGEAL--SVAAAKRLKLLLFAPVACTSLE--YNIRVYCLHDTHDALKEVVQLEKQLGGQL	692
		: : : : : : : : : :	
Db	655	AGHPRRNSLSAAKRVHLAVFGPTMSAYRRPFELRVYCVPETGAAMESVWKQED--GSRL	712
Qy	693	IQEPR--VLHFKDSYHNLRSLIHDV-PSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFT	749

A;Description: required for guidance of pioneering axons and cells migrating dorsally along the body wall; proposed to be a receptor on the surface of the motile cells

C;Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology; thrombospondin type 1 repeat homology

C;Keywords: alternative splicing; duplication; glycoprotein; receptor; transmembrane protein

F;30-947/Product: unc-5 protein, short form #status predicted <ALT>

F;46-116/Domain: immunoglobulin homology <IM1>

F;153-211/Domain: immunoglobulin homology <IM2>

F;229-300/Domain: thrombospondin type 1 repeat homology #status atypical <THR1>

F;301-354/Domain: thrombospondin type 1 repeat homology <THR2>

F;365-390/Domain: transmembrane #status predicted <TMM>

F;512-559/Domain: SH3 homology <SH3>

F;53-114,65-112,160-209/Disulfide bonds: #status predicted

F;206/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.4%; Score 977; DB 1; Length 947;
Best Local Similarity 28.7%; Pred. No. 2.4e-62;
Matches 265; Conservative 168; Mismatches 379; Indels 110; Gaps 31;

```
Qy      49 EPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIER--STDGSSGLPTMEVRI 106
          :: |::|::| |::| |::| ::| |::| |::| ::|
Db      37 QPKSGYVIRNKPLRLQCRANHATKIRYKCSSKWID--DSRIEKLIGTDSTSGVGIDASV 94

Qy     107 NVSRQQVEKVFGLLEEYWCQCVAWSSSG-----TTKSQKAYIRIARLRKNFEQEPLAKEVS 161
          ::|| |:: |::| |::| |::| |::| |::| |::|
Db      95 DISRIDVDTSGHVDAFQCQCYA---SGDDDQDVVASDVATVHLAYMRKHFLKSPVAQRVQ 151

Qy     162 LEQGIVLPCRPPPEGIPPAEVEWLRNEDLVDPSPDNVYITREHSLVVRQARLADTANYTC 221
          : ||:: || |::| ::| |::| |::| |::| |::|
Db     152 EGTTLQLPCQAPESDPKAELTWYKDGVVVQP--DANVIRASDGLSIMSAARLSDSGNYTC 209

Qy     222 VAKNIVARRRSASAAVIVYVNGGWSTWTEW-SVCSASCG-----RGWQKR 265
          | |:: |:: |::|::| |::| |::| |::|
Db     210 EATNVANSRKTDPEVQIYVDGGWSESWPWIGTCHVDCPLLQHAHRIRDPHDVLPHQRR 269

Qy     266 SRSCNPNAPLNGGAFCEGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWRSDRECS 325
          ::|| ||::| |::| ::| |::| ||::| ||::| ||::|
Db     270 TRTCNNPAPLNDGEYCKGEEEMTRSKVPCKLDGGWSSWSDWSACSSSCHRYRTRACTVP 329

Qy     326 APRNGGEECQGTDLDTNRCTSDLCVHSASG--PEDVALYVGLIAVAVCLVLLLVLILVY 383
          | ||:: | | || |::| |::| |::| |::| |::|
Db     330 PPMNGGQPCFGDDLMTQECPAQLCTADSSRIVISDTAVYGSVASIFIVASFILAILAMFC 389

Qy     384 CR-----KKEGLDSDVADSSSILTSQFQVPSIKPSKADNPHLLTI----- 422
          |:: |::|::| |::| |::| |::| |::|
Db     390 CKRGNSKSKPLKPQKMNSEKAGGIYYS---EPPGVRRLLLEHQHGTLLGEKISSCSQYF 446

Qy     423 -QPDLTSTTT-----TYQGSCLPRQDGSPKQFQLTNGHLLSPLGGGRHTLHHSSPT-SE 474
          | |::| |::| |::| |::| |::| |::|
Db     447 EPPPLPHSTTLRSGKSAFSGYSSTRNAGSRAALIQCSSSSSGSGGKRTMLRTSSSNCS 506

Qy     475 AEEFVSRLSTQNYFRSLPRGTS-NMTYGTNFGRLMIPNTGISLLIPPDAIPRGIYE 533
          :: |:: |::| |::| |::| |::| |::|
Db     507 DDNYATLYDYMEDKSVLGLDTSQNIVAAQIDSNGARLSLSKSGARLIVPELAVEGEKM-- 564
```

```

Qy      534 IYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGV-----LLTRPVILAMDHCGEPS-D 585
      :|| :      : |      :: |||:: | |      :| |||::: ||      | |
Db      565 LYLAVSDTLTDQPHLKPIESALSPVIVIGQCDVSMSAHDNILRRPVVVSFRHCASTFPRD 624

Qy      586 SWSLRLKKQSCEGS-WEDVLHLGEEAPSHLYYCQLEASA-----CYVFTEQLGRFAL 636
      :| | : ||| | : : :||| : : | |      | :| | | | |
Db      625 NWQFTL--YADEGSGWQKAVTIGEENLNTNMFVQFEQPGKKNDGFGWCHVMTYSLARLML 682

Qy      637 VGEAL--SVAAAKRLKLLLFAPVACTSLE--YNIRVYCLHDTHDALKEVVQLEKQLGGQL 692
      |      |::||| : | :| | : :      : :||| : :| | : | : | :|
Db      683 AGHPRRNSLSAAKRVHLAVFGPT EMSAYRRPFELRVYCV PETGAAMESVWKQED--GSRL 740

Qy      693 IQEPR--VLHFKDSYHNLRLSIHDV-PSSLWKS KLLVSYQEIPFYHIWNGTQRYLHCTFT 749
      : |      :| : | || : | || |      :| | :      | ||| :
Db      741 LCESNDFILNEKG---NLCICIEDVIPGFS CDGPEVVEISETQHRFV---AQNLHCSLK 794

Qy      750 LERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAF 809
      : | : : :| :| : : : : : : :| | | : |
Db      795 FRPKEINGSQFSTRVIVYQKASSTEPMMV--EVSNEPELYDATSEEREKGSVCV----EF 848

Qy      810 KIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKP--SPTAMILNLWEARH 867
      ::|| : : : || | :||| ||:||| | :| |||| | |||:::|:||||
Db      849 RLPFGVKDELARLLDMPNESHSDWRGLAKKLHYDRYLQFFASFPDCSPTSLLLDLWEASS 908

Qy      868 FPNGN-LSQLAAAVAGLGQPD A 888
      : : | : :|:||||
Db      909 SGSARAVPDL LQTLRVMGRPDA 930

```

RESULT 3

T00026

brain-specific angiogenesis inhibitor 1 - human

N;Alternate names: BAI1 protein

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004

C;Accession: T00026

R;Nishimori, H.; Shiratsuchi, T.; Urano, T.; Kimura, Y.; Kiyono, K.; Tatsumi, K.; Yoshida, S.; Ono, M.; Kuwano, M.; Nakamura, Y.

submitted to the EMBL Data Library, June 1997

A;Reference number: Z14064

A;Accession: T00026

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1584 <NIS>

A;Cross-references: UNIPROT:O14514; EMBL:AB005297; NID:d1175078; PID:d1024528

A;Experimental source: brain

C;Genetics:

A;Gene: GDB:BAI1

A;Cross-references: GDB:9838088; OMIM:602682

A;Map position: 8q24-8q24

F;408-462/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 6.2%; Score 298.5; DB 2; Length 1584;

Best Local Similarity 33.5%; Pred. No. 5.2e-13;

Matches 78; Conservative 35; Mismatches 91; Indels 29; Gaps 11;

```

Qy      124 CQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEW 183

```


F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>
 F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
 F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
 F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
 F;553-588/Domain: EGF homology <EGF1>
 F;652-691/Domain: EGF homology <EGF>
 F;928-930/Region: cell attachment (R-G-D) motif
 F;151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent)
 #status predicted
 F;167-226/Disulfide bonds: #status predicted
 F;266,270/Disulfide bonds: interchain #status predicted
 F;612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 6.2%; Score 296.5; DB 1; Length 1172;
 Best Local Similarity 30.5%; Pred. No. 4.9e-13;
 Matches 78; Conservative 28; Mismatches 105; Indels 45; Gaps 9;

```

Qy      209 RQARLADTANYTCVAKNIVARRRSASAA-VIVYVNGGWSTWTEWSVCSASCGRGWQKRSR 267
          :: | | : ||: :| | | | : :||| |: || || :|| | | |
Db      403 QRGRSCDVTSTNTCLGPSIQTRACSLSKCDTRIRQDGGWSHWPWSSCSVTCGVGNITRIR 462

Qy      268 SCTNPAPLNGGAFCEGQNQKTAC-ATLCPVDGWSWPWSKWSACGLDCT---HWRSRECS 323
          | :| | || |:| : | || ||:| | |||| | || | : | | :| | :
Db      463 LCNSPVPQMGGKNCKGSGRETKACQGAPCPIDGRWSPWSPWSACTVTCAGGIRERTRVCN 522

Qy      324 DPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVY 383
          | |: ||: | | : : | | | | | | | | | |
Db      523 SPEPQYGGKACVGDVQERQMCNKRSC-----PVDGCLSNPCFPGAQC----- 564

Qy      384 CRKKEGLDSDVADSSILTSGFQPVSI--KPSKADNPHLLTIQPDLSSTTTT-----TYQ 434
          | | | : || || : : : : : ||: :|: |
Db      565 -----SSFPDGS-WSCGFPCVGFILGNTHCEDLDECALVPDICFSTSKVPRCVNTQP 615

Qy      435 GSLC----PRQDGPSP 446
          | | || | |
Db      616 GFHCLPCPPRYRGNQP 631
  
```

RESULT 5

JC5928

semaphorin F precursor - human

C;Species: Homo sapiens (man)

C;Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

C;Accession: JC5928

R;Simmons, A.D.; Pueschel, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.
 Biochem. Biophys. Res. Commun. 242, 685-691, 1998

A;Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-
 chat candidate interval.

A;Reference number: JC5928; MUID:98125554; PMID:9464278

A;Accession: JC5928

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-1074 <SIM>

A;Cross-references: UNIPROT:Q13591; GB:U52840; NID:g2772583; PIDN:AAC09473.1;
 PID:g2772584

A;Experimental source: brain

C;Comment: This protein disrupts normal brain development and leads to some of the features of Cri-du-chat.

C;Genetics:

A;Gene: sema4

C;Superfamily: human semaphorin F; thrombospondin type 1 repeat homology

F;1-20/Domain: signal sequence #status predicted <SIG>

F;50-533/Domain: semaphorin #status predicted <SEM>

F;840-896/Domain: thrombospondin type 1 repeat homology <THR3>

F;971-993/Domain: transmembrane #status predicted <TMM>

Query Match 6.1%; Score 293; DB 2; Length 1074;

Best Local Similarity 45.8%; Pred. No. 7.8e-13;

Matches 54; Conservative 11; Mismatches 49; Indels 4; Gaps 2;

```
Qy      241 VNGGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATL-CPVDG 299
      ||| || || || || | || : | | | | | || | | | | | | | | |
Db      783 VNGAWSAWTSWSQCSRDCSRGIRNRKRVCNNPEPKYGGMPCLGPSLEYQECNTLPCPVDG 842

Qy      300 SWSPWSKWSACGLDC---THWRSRECSDPAPRNGGEECQGTDLDTNRNCTSDLCVHSAS 354
      || || | : | | : | : | | | | | | | | | : | : | | |
Db      843 VWSCWSPWTKSATCGGGHYMRTRSCSNPAPAYGGDICLGLHTEELCNTQPCPESWS 900
```

RESULT 6

A42587

thrombospondin 2 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C;Accession: A42587; A39851

R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992

A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell growth and development.

A;Reference number: A42587; MUID:92147683; PMID:1371115

A;Accession: A42587

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: nucleic acid

A;Residues: 1-1172 <LAH>

A;Cross-references: UNIPROT:Q03350; GB:L07803; GB:M87275; NID:g340421;
PIDN:AAA53064.1; PID:g567241

A;Note: sequence extracted from NCBI backbone (NCBIP:81502)

R;Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.;
Dixit, V.M.

J. Biol. Chem. 266, 12821-12824, 1991

A;Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.

A;Reference number: A39851; MUID:91302287; PMID:1712771

A;Accession: A39851

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-873 <BOR>

A;Cross-references: GB:M64866; NID:g201994; PIDN:AAA40432.1; PID:g201995

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

C;Keywords: calcium binding; glycoprotein

F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>

F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>

F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F;553-588/Domain: EGF homology <EGF1>
F;652-691/Domain: EGF homology <EGF>

Query Match 6.1%; Score 293; DB 2; Length 1172;
Best Local Similarity 38.0%; Pred. No. 8.7e-13;
Matches 60; Conservative 22; Mismatches 66; Indels 10; Gaps 5;

```
Qy      209 RQARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSR 267
          :: | | : ||: :| | | : |||| | : || || :|| | | |
Db      403 QGRGRSCDVTSNTCLGPSIQTRTCSLGKCDTRIRQNGGWSHWPWSSCSVTCTGCVGNVTRIR 462

Qy      268 SCTNPAPLNGGAFCEGQNVQKTAC-ATLCPVDGSWSPWSKWSACGLDCT---HWSRECS 323
          | :| | || |:| : | ||:| | |||| |||| : | || | :
Db      463 LCNSPVPQMGGKNCKGSGRETKPCQRPIDGRWSPWSPWSACTVTCAGGIRERSRVCN 522

Qy      324 DPAPRNGGEECQG---TD---LDTRNCTSDLCVHSASGP 356
          | |: ||::| | |: : |:| | |: : |
Db      523 SPEPQYGGKDCVGDVTEHQMCNKRSCPIDGCLSNPCFP 560
```

RESULT 7

T18856

angiogenesis inhibitor homolog - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T18856; T24653

R;McMurray, A.

submitted to the EMBL Data Library, July 1995

A;Reference number: Z19031

A;Accession: T18856

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1444 <WIL>

A;Cross-references: UNIPROT:Q8MYA8; EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028;
CESP:C02B4.1

A;Experimental source: clone C02B4

R;McMurray, A.

submitted to the EMBL Data Library, July 1995

A;Reference number: Z19917

A;Accession: T24653

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1444 <WI2>

A;Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1

A;Experimental source: clone T07C5

C;Genetics:

A;Gene: CESP:C02B4.1

A;Map position: X

A;Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3;
509/3; 566/2; 625/1; 696/2; 786/3; 812/2; 878/3; 971/1; 1007/3; 1067/1; 1099/3;
1180/3; 1273/2; 1305/1; 1363/1; 1388/2

Query Match 5.8%; Score 276; DB 2; Length 1444;
Best Local Similarity 27.2%; Pred. No. 2e-11;
Matches 73; Conservative 28; Mismatches 97; Indels 70; Gaps 12;

```

Qy      123 WCQCVAWSSSGTTKSQKAYIRIARLRKNFEQ-----EPLAKEVSLEQGIVLPCRPEGI 176
      | : :||:      : : | : | :      :      ||| :|| |
Db      1134 WSEWSSWSAC-----SCFSLTSTRRRFCQVVDPTVQGFCAGAILEQ---IPCAPGSCS 1183

Qy      177 PPAE-----VEW-----LRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN 218
      | |      ||      :|| :|      :      | | :
Db      1184 PSAGGWSLWSEWSSCSKDCGDTGHQIRNRMCSPE-----IPSNRGAYCSG 1228

Qy      219 YT-----CVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPA 273
      |:      || |: : :      |:|||: || || |: | | : |:| | ||
Db      1229 YSFDQRPCVMDNVCSDEK-----VDGGWTDWTAWSECTDYCRNGHRSRTRFCANPK 1279

Qy      274 PLNGGAFCEGQNVQKTAC--ATLCPV-DGSWSPWSKWSACGLDC---THWRSRECSDPAP 327
      | ||| | | : : |      | : || || || |: | |      | ||| | |
Db      1280 PSQGAQCTGSDFELNPCFDPARCHLRDGGWSTWSDWTPCSASCGFGVQTRDRSCSSPEP 1339

Qy      328 RNGGEECQGTDLDTNRCTSDLCVHSASG 355
      : ||: | | | | | : |
Db      1340 K-GGQSCSGLAHQTSLCDLPACDHESDG 1366

```

RESULT 8

T00326

hypothetical protein KIAA0550 - human

C;Species: Homo sapiens (man)

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 15-Mar-2004

C;Accession: T00326

R;Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.

DNA Res. 5, 31-39, 1998

A;Title: Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.

A;Reference number: Z14086; MUID:98290545; PMID:9628581

A;Accession: T00326

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-984 <NAG>

A;Cross-references: EMBL:AB011122; NID:g3043623; PIDN:BAA25476.1; PID:g3043624

A;Experimental source: brain

C;Genetics:

A;Note: KIAA0550

F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 5.7%; Score 275; DB 2; Length 984;

Best Local Similarity 39.0%; Pred. No. 1.4e-11;

Matches 57; Conservative 20; Mismatches 53; Indels 16; Gaps 6;

```

Qy      220 TCVA-----KNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPA 273
      |||:      : | : : |:| | |: ||:| | :||| : |:| | |
Db      317 TCVSPYGTHCSGPLRESRVCNNALCPVHGVWEEWSPWSLCSFTCGRGQRTRTRSCT--P 374

Qy      274 PLNGGAFCEGQNVQKTAC-ATLCPVDGSWSPWSKWSACGLDC---THWRSRECSDPAPRN 329
      | || |||      | ||||| | || || | : | | |||:| : | :
Db      375 PQYGGRPCEGPETHHKPCNIALCPVDGQWQEWSSWSQCSVTCSNGTQQRSRQCT--AAAH 432

```

Qy 330 GGEECQGTDLDTNRCTSDLCVHSASG 355
 || ||:| :| | : | :|:|
 Db 433 GGSECRGPWAESRECYNPEC--TANG 456

RESULT 9

T00028

brain-specific angiogenesis inhibitor 3 - human

N;Alternate names: BAI3 protein

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004

C;Accession: T00028

R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.

Cytogenet. Cell Genet. 79, 103-108, 1997

A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-specific angiogenesis inhibitor 1 (BAI 1).

A;Reference number: Z14066; MUID:98194217; PMID:9533023

A;Accession: T00028

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1522 <SHI>

A;Cross-references: UNIPROT:O60242; EMBL:AB005299; NID:g3021700;

PIDN:BAA25363.1; PID:g3021701

A;Experimental source: brain

C;Genetics:

A;Gene: GDB:BAI3

A;Cross-references: GDB:9838090; OMIM:602684

A;Map position: 6q12-6q12

F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 5.7%; Score 275; DB 2; Length 1522;
 Best Local Similarity 39.0%; Pred. No. 2.5e-11;
 Matches 57; Conservative 20; Mismatches 53; Indels 16; Gaps 6;

Qy 220 TCVA-----KNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCNPA 273
 |||: : | : |:| | |: ||: || : |||| : |: ||||
 Db 317 TCVSPYGTHCSGPLRESRVCNNTALCPVHGVEEWSPWSLCSFTCGRGQRTTRTRSC--P 374
 Qy 274 PLNGGAFCEGQNVQKTAC-ATLCPVDGWSWPWSKWSACGLDC---THWRSRECSDPAPRN 329
 | ||. ||| | ||||| | || | : | | |||:| : | :
 Db 375 PQYGGRPCEGPETHHKPCNIALCPVDGQWQEWSSWSQCSVTCSNGTQQRSRQCT--AAAH 432
 Qy 330 GGEECQGTDLDTNRCTSDLCVHSASG 355
 || ||:| :| | : | :|:|
 Db 433 GGSECRGPWAESRECYNPEC--TANG 456

RESULT 10

T00027

brain-specific angiogenesis inhibitor 2 - human

N;Alternate names: BAI2 protein

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004

C;Accession: T00027

R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.

Cytogenet. Cell Genet. 79, 103-108, 1997

A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-specific angiogenesis inhibitor 1 (BAI 1).
 A;Reference number: Z14066; MUID:98194217; PMID:9533023
 A;Accession: T00027
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-1572 <SHI>
 A;Cross-references: UNIPROT:O60241; EMBL:AB005298; NID:g3021698;
 PIDN:BAA25362.1; PID:g3021699
 A;Experimental source: brain
 C;Genetics:
 A;Gene: GDB:BAI2
 A;Cross-references: GDB:9838089; OMIM:602683
 A;Map position: 1p35-1p35

Query Match 5.7%; Score 274.5; DB 2; Length 1572;
 Best Local Similarity 19.2%; Pred. No. 2.8e-11;
 Matches 176; Conservative 108; Mismatches 307; Indels 327; Gaps 38;

Qy	173	PEGIPPAEVEWLRNEDLVDP	SLDPNVY-----ITREHSLVVRQARL	213
		: : :	:	
Db	271	PEEEPKVKTQWPRSAD-----	EPGLYMAQTGDPAAEWSPWSVCSLTCGQGLQVR-TRS	323
Qy	214	ADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSC	TNP	273
		:: : : : : : : :		
Db	324	CVSSPYGTLCSGPLRETRPCNNSATCPVHGVWEEWGSWSLCSRSCGRGSRSRMRTC	V--P	381
Qy	274	PLNGGAFCEGQNVQKTACA-TLCPVDGWSWSPWSKWSACGLDC---	THWRSRECS	328
		: : : :		
Db	382	PQHGGKACEGPELQTKLCSMAACPVEGQWLEWGPWGPCSTSCANGTQQR	SRKCSVAGPAW	441
Qy	329	-----	NGGEECQ	335
			:	
Db	442	ATCTGALTDTRECSNLECPATDSKWGPWNAWSLCSKTCDTGWQRRFRMCQATGTQGY	PCE	501
Qy	336	GTDLDRNCTSDLC--VHSASGPEDVAL-----		361
		: : :		
Db	502	GTGEEVKPCSEKRCPAFHEMCRDEYVMLMTWKAAAAGEIIYNKCPPNASGSASRRCLLSA		561
Qy	362	----YVGLIAVAVCL---VLLLLVLILVYCRKKEGLSDVADSSILTS	SGFQPVSIKPSKA	414
		: : : : : : : : : : : :		
Db	562	QGVAYWGLPSFARCISHEYRYLYLSLREHLAKGQRLAGEGMSQVVRS-LQELLAR	RTYY	620
Qy	415	DNPHLLTIQPDLS	TTTTTYQGS	461
		: : : : : :	: :	
Db	621	SGDLLFSVDILRNVTDTFKRATYVPSADDVQRFQVVSFMDAENKEKWDDAQQVSP--	G	678
Qy	462	GRHTLHHSSPTSEAEFV-----	SRLSTQNYFRSLPRG-----TSNMTYGT	504
		: : : : : : :		
Db	679	SVHLLR-----VVEDFIHLVGDAKAFQSS	LIVTDNLVISIQREPVS	732
Qy	505	FLG-----	GRLMIPNTGISLLIP-----PDAIPRGK-----	530
		: : :		
Db	733	RRGMKDWVRHSEDRFLPK	EVLSSPGK	792
Qy	531	-----IYE-IYLT	LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVIL	574
		: : : : : :		

Db 793 PADPDESSYFVIGAVLYRTLGLILPPP---RPPLAVTSRVMT--VTVRPPTQPPAEPLIT 847

Qy 575 A-----MDHCGEPPSPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQ-LEASACYV-- 626
 :: :| || : : | | : || || | :

Db 848 VELSYIINGTTDPHCASWDYS-RADASSGDWD-----TENCQTLETQAAHTRC 894

Qy 627 FTEQLGRFALVGE-----ALSVAANKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEV 681
 : | ||:: : | : : : |:: | : | : : | |

Db 895 QCQHLSTFAVLAQPPKDLTLELAGSPSVPLVIGCAVSCMALLTLAIYA-----AFWRF 948

Qy 682 VQLEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDPSSLWKSLLVSYQEIPFYHIWNGTQ 741
 :: | : : | ||| : |:: : || : : |

Db 949 IKSERSI-----ILLNFCLSI--LASNI---LILVGQSRVLSKGVCTMTA 988

Qy 742 RYLHCTFTLERVSPSTSDLACKLWV-----WQVEGDG 773
 : || | | : || | | : |

Db 989 AFLHFFF-----LSSFCEWVLTEAWQSYLAVIGRMTRLVRKRFLCLGWGLPALV 1037

Qy 774 QSFSINFNITKDTRFAELLALESEAG-VPALVGPSA-----FKIPFLIRQKI-----IS 821
 : | : | || : | | | : | ||| : | | : : | : ||

Db 1038 VAVSVGFTRTKGYGTSSYCWLSLEGGLLYAFVGPAAVIVLVNMLIGIIVFNKLMARDGIS 1097

Qy 822 SLDPPCRRGAD---WRTL 836
 | |:: | : |

Db 1098 DSKKKQRAGSERC PWASL 1115

RESULT 11

A40558

thrombospondin 1 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004

C;Accession: A40558; A37905; B42587; S68787

R;Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.

Genomics 11, 587-600, 1991

A;Title: Characterization of the murine thrombospondin gene.

A;Reference number: A40558; MUID:92128941; PMID:1774063

A;Accession: A40558

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1170 <LAW>

A;Cross-references: UNIPROT:P35441; GB:M62449; GB:M62450; GB:M62451; GB:M62452;

GB:M62453; GB:M62454; GB:M62455; GB:M62456; GB:M62457; GB:M62458; GB:M62459;

GB:M62460; GB:M62461; GB:M62462; GB:M62463; GB:M62464; GB:M62465; GB:M62466;

GB:M62467; GB:M62468; GB:M62469; GB:M62470; NID:g511867; PIDN:AAA50611.1;

PID:g511869

R;Bornstein, P.; Alfi, D.; Devarayalu, S.; Framson, P.; Li, P.

J. Biol. Chem. 265, 16691-16698, 1990

A;Title: Characterization of the mouse thrombospondin gene and evaluation of the role of the first intron in human gene expression.

A;Reference number: A37905; MUID:90375546; PMID:2398070

A;Accession: A37905

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-490 <BOR>

A;Cross-references: GB:J05605; GB:J05606; NID:g201991; PIDN:AAA40431.1;
 PID:g554390
 R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
 J. Biol. Chem. 267, 3274-3281, 1992
 A;Title: Characterization of mouse thrombospondin 2 sequence and expression
 during cell growth and development.
 A;Reference number: A42587; MUID:92147683; PMID:1371115
 A;Accession: B42587
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-1152,'P',1154-1170 <LAH>
 A;Cross-references: GB:M87276
 A;Note: sequence extracted from NCBI backbone (NCBIP:81501)
 R;Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F.
 FEBS Lett. 387, 36-41, 1996
 A;Title: Expression and initial characterization of recombinant mouse
 thrombospondin 1 and thrombospondin 3.
 A;Reference number: S68787; MUID:96234006; PMID:8654563
 A;Accession: S68787
 A;Molecule type: protein
 A;Residues: 19-26,'X',28-37 <CHE>
 C;Complex: homotrimer, disulfide linked
 C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat
 homology; von Willebrand factor type C repeat homology
 C;Keywords: calcium binding; glycoprotein; homotrimer
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
 F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
 F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
 F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
 F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
 F;551-586/Domain: EGF homology <EGF>
 F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.6%; Score 270.5; DB 2; Length 1170;
 Best Local Similarity 32.2%; Pred. No. 3.7e-11;
 Matches 57; Conservative 24; Mismatches 71; Indels 25; Gaps 5;

```

Qy      207 VVRQARLADTANYTCVAKNIVAR-----RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
          : :: | | : | | : : | : | : ||| | : || || : |
Db      399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-----DGGWSHWSPWSSCSVTC 451

Qy      259 GRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTAC-ATLCPVDGWSWPWSKWSACGLDC--- 314
          | | | | | : | | | | : : | | | : | | | | | : |
Db      452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKKDACPINGGWGPWPWDICSVTCGGG 511

Qy      315 THWRSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVALYVGLIAVAVC 371
          ||| | : | | : | : | | : : | | | | | | | |
Db      512 VQRRSRLCNNPTPQFGGKDCVGDVTENQVCNKQDC-----PIDGCLSNPCFAGAKC 562

```

RESULT 12

TSHUP1

thrombospondin 1 precursor - human

C;Species: Homo sapiens (man)

C;Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C;Accession: A26155; A34274; A30140; A25812; A05172; A42927

R;Lawler, J.; Hynes, R.O.
 J. Cell Biol. 103, 1635-1648, 1986
 A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.
 A;Reference number: A26155; MUID:87057617; PMID:2430973
 A;Accession: A26155
 A;Molecule type: mRNA
 A;Residues: 1-1170 <LAW>
 A;Cross-references: UNIPROT:P07996; GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing
 R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.
 J. Biol. Chem. 264, 11222-11227, 1989
 A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.
 A;Reference number: A34274; MUID:89291870; PMID:2544587
 A;Accession: A34274
 A;Molecule type: DNA
 A;Residues: 1-166 <LAH>
 A;Cross-references: GB:J04835
 R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Frazier, W.A.
 J. Cell Biol. 108, 729-736, 1989
 A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.
 A;Reference number: A30140; MUID:89139590; PMID:2918029
 A;Accession: A30140
 A;Molecule type: mRNA
 A;Residues: 1-83, 'A', 85-522, 'A', 524-1170 <HEN>
 A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing
 R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.
 Biochemistry 25, 8418-8425, 1986
 A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.
 A;Reference number: A25812; MUID:87157592; PMID:3030396
 A;Accession: A25812
 A;Molecule type: mRNA
 A;Residues: 1-83, 'A', 85-397 <KOB>
 A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354
 R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.
 Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
 A;Reference number: A05172; MUID:86287276; PMID:3461443
 A;Accession: A05172
 A;Molecule type: mRNA
 A;Residues: 1-83, 'A', 85-374, 'RC' <DIX>
 A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing
 R;Sun, X.; Skorstengaard, K.; Mosher, D.F.
 J. Cell Biol. 118, 693-701, 1992
 A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.
 A;Reference number: A42927; MUID:92348511; PMID:1379247
 A;Accession: A42927

A;Molecule type: protein
 A;Residues: 987-1003 <SUN>
 A;Note: Cys-992 is shown to have a free sulfhydryl
 C;Genetics:
 A;Gene: GDB:THBS1; TSP1; TSP
 A;Cross-references: GDB:120438; OMIM:188060
 A;Map position: 15q15-15q15
 A;Introns: 23/1
 A;Note: the list of introns may be incomplete
 C;Complex: homotrimer, disulfide linked
 C;Function:
 A;Description: participates in cell migration and adhesion, and in platelet aggregation
 C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology
 C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
 F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
 F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
 F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
 F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
 F;551-586/Domain: EGF homology <EGF1>
 F;650-689/Domain: EGF homology <EGF2>
 F;926-928/Region: cell attachment (R-G-D) motif
 F;171-232/Disulfide bonds: #status predicted
 F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;270,274/Disulfide bonds: interchain #status predicted
 F;610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
 F;1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 5.6%; Score 268.5; DB 1; Length 1170;
 Best Local Similarity 32.9%; Pred. No. 5.1e-11;
 Matches 51; Conservative 24; Mismatches 61; Indels 19; Gaps 4;

```

Qy      207 VVRQARLADTANYTCVAKNIVAR-----RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
      : : | | : | | : : | : | : | | | | : | | | | : |
Db      399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-----DGGWSHWSPWSSCSVTC 451

Qy      259 GRGWQKRSRSCCTNPAPLNGGAFCEGQNVQKTAC-ATLCPVDGSWSPWSKWSACGLDC--- 314
      | | | | | : : | | | | : : | | | : : | | | | | : |
Db      452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKKDACPINGGWGPWSPWDICSVTCGGG 511

Qy      315 THWRSRECSDEAPRNGGEECQGTDLDTNRNCTSDL 349
      . | | | | : : | | : | | : : | | : : | |
Db      512 VQKRSRLCNNPTPQFGGKDCVGDVTENQICNKQDC 546
  
```

RESULT 13

A39804

thrombospondin precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C;Accession: A39804

R;Lawler, J.; Duquette, M.; Ferro, P.

J. Biol. Chem. 266, 8039-8043, 1991

A;Title: Cloning and sequencing of chicken thrombospondin.
A;Reference number: A39804; MUID:91217026; PMID:2022631
A;Accession: A39804
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 <LAW>
A;Cross-references: UNIPROT:P35440; GB:M60853; NID:g212763; PIDN:AAA51437.1; PID:g212764
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology
F;325-383/Domain: von Willebrand factor type C repeat homology <VWC>
F;386-437/Domain: thrombospondin type 1 repeat homology <THR1>
F;442-498/Domain: thrombospondin type 1 repeat homology <THR2>
F;499-555/Domain: thrombospondin type 1 repeat homology <THR3>
F;658-697/Domain: EGF homology <EGF>

Query Match 5.5%; Score 263; DB 1; Length 1178;
Best Local Similarity 36.2%; Pred. No. 1.3e-10;
Matches 58; Conservative 16; Mismatches 70; Indels 16; Gaps 5;

Qy 210 QARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGGWSTWTEWSVCASCGRGWQKRSRS 268
: | | | : | | : : || | : | | | : | | |
Db 410 RGRSCDVTRSACTGPHIQTRMCSFKKCDHRIRQDGGWSHWSPWSSCSVTGCVGNITRIRL 469
Qy 269 CTNPAPLNGGAFCEGQNVQKTACATL-CPVDGSWSPWSKWSACGLDC---THWRSRECS 324
| : | | | | : | | | : | | | | : | | | :
Db 470 CNSPIPQMGGKNCVGNGRETEKCEKAPCPVNGQWGPWPWSACTVTCGGGIRERSRLCNS 529
Qy 325 PAPRNGGEECQGTDLDT-----RNCTSDLCVHSASGP 356
| | : | | | | : | | | : : |
Db 530 PEPQYGGKPCVG---DTKQHDMCNKRDCPIDGCLSNPCFP 566

RESULT 14

S29126

properdin precursor [validated] - human

N;Alternate names: factor P

C;Species: Homo sapiens (man)

C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004

C;Accession: S29126; S16150; A05319; T45112; T45113

R;Nolan, K.F.; Kaluz, S.; Higgins, J.M.G.; Goundis, D.; Reid, K.B.M.
Biochem. J. 287, 291-297, 1992

A;Title: Characterization of the human properdin gene.

A;Reference number: S29126; MUID:93038568; PMID:1417780

A;Accession: S29126

A;Molecule type: DNA

A;Residues: 1-469 <NOL1>

A;Cross-references: UNIPROT:P27918; EMBL:X70872; NID:g35679; PIDN:CAA50220.1; PID:g35680

R;Nolan, K.F.; Schwaeble, W.; Kaluz, S.; Dierich, M.P.; Reid, K.B.M.
Eur. J. Immunol. 21, 771-776, 1991

A;Title: Molecular cloning of the cDNA coding for properdin, a positive regulator of the alternative pathway of human complement.

A;Reference number: S16150; MUID:91184288; PMID:2009915

A;Accession: S16150

A;Molecule type: mRNA

A;Residues: 1-456, 'R', 458-469 <NOL2>

A;Cross-references: EMBL:X57748
 R;Reid, K.B.M.; Gagnon, J.
 Mol. Immunol. 18, 949-959, 1981
 A;Reference number: A05319; MUID:82195224; PMID:7341961
 A;Accession: A05319
 A;Molecule type: protein
 A;Residues: 28-53,'Q',55-59,'G',61,'I',63;137-138,'P',140-141,'P',143-144,'X',146-148,'Y',150,'S',152,'Y',154-156,'XSXGXA';162-163,'E',165-172,'X',174-176,'X',178,'V',180;223-228,'X',230-232,'GX',235-238,'GH',241-245;248-251,'X',253-257,'P',259,'G',261,'XPP',265-266,'X',268-269;280-285,'X',287-290,'X',292,'H',294-300,'SXXX',305-307,'X',309-315,'K',317;333-341,343-357,'X',359-362,'EXE';393-404,'QK',407;421-427,'R',429-443,'TKV',447-448,'XX',451,'RX',454-455 <REI>
 R;Westberg, J.; Nordin-Fredrikson, G.; Truedsson, L.; Sjöholm, A.G.; Uhlen, M. submitted to the EMBL Data Library, May 1997
 A;Reference number: Z22914
 A;Accession: T45112
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-54,'X',56-73,'X',75-99,'W',101-469 <WES1>
 A;Cross-references: EMBL:AF005665; PIDN:AAB63280.1
 A;Experimental source: genomic DNA from individual with properdin deficiency type II
 A;Accession: T45113
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-60,'X',62-413,'D',415-452,'XX',455-469 <WE2>
 A;Cross-references: EMBL:AF005666; PIDN:AAC51626.1
 A;Experimental source: genomic DNA from individual with properdin deficiency type III
 R;Hartmann, S.; Hofsteenge, J.
 J. Biol. Chem. 275, 28569-28574, 2000
 A;Title: Properdin, the positive regulator of complement, is highly C-mannosylated.
 A;Reference number: A59360; MUID:20435812; PMID:10878002
 A;Contents: annotation
 A;Note: identification and location of C-mannosylation sites by mass-spectroscopy
 C;Genetics:
 A;Gene: GDB:PFC
 A;Cross-references: GDB:120275; OMIM:312060
 A;Map position: Xp11.3-Xp11.23
 A;Introns: 26/1; 76/2; 135/1; 192/1; 256/1; 314/1; 378/1; 415/2
 C;Complex: a mixture of homodimers, homotrimers and homotetramers
 C;Function:
 A;Description: protects C3 convertase (C3bBb) from rapid inactivation
 A;Pathway: complement alternate pathway
 C;Superfamily: human properdin precursor; thrombospondin type 1 repeat homology
 C;Keywords: complement alternate pathway; glycoprotein; homodimer; homotetramer; homotrimer; plasma
 F;1-27/Domain: signal sequence #status predicted <SIG>
 F;28-469/Product: properdin #status experimental <MAT>
 F;76-128/Domain: thrombospondin type 1 repeat homology <THR1>
 F;135-191/Domain: thrombospondin type 1 repeat homology <THR2>
 F;192-255/Domain: thrombospondin type 1 repeat homology <THR3>
 F;256-313/Domain: thrombospondin type 1 repeat homology <THR4>
 F;314-377/Domain: thrombospondin type 1 repeat homology <THR5>

F;378-440/Domain: thrombospondin type 1 repeat homology <THR6>
F;83,86,139,142,145,196,199,260,263,321,324,382,385,388/Modified site: 2'-
mannosyl-tryptophan (Trp) #status experimental
F;428/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.1%; Score 243; DB 1; Length 469;
Best Local Similarity 39.5%; Pred. No. 1.1e-09;
Matches 45; Conservative 14; Mismatches 43; Indels 12; Gaps 4;

```
Qy      243 GGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACAT--LCPVDGS 300
      |||| | | || :| :| : | |:| :|| || | || : || | :|| | :
Db      137 GGWSGWGPWEPCSVTCSKGRTRRRACNHPAPKCGG-HCPGQAQEQSEACDTQQVCPTHGA 195

Qy      301 WSPWSKWSACGLDC-----THWRSRECSDPAP--RNGGEECQGTDLDTNRCT 345
      |: | |: | | | ||:| | | : |: | | : | ||
Db      196 WATWGPWTPCSASCHGGPHEPKETRSRKCSAPEPSQKPPGKPCPLAYEQRRCT 249
```

RESULT 15

S05478

properdin - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004

C;Accession: S05478

R;Goundis, D.; Reid, K.B.M.

Nature 335, 82-85, 1988

A;Title: Properdin, the terminal complement components, thrombospondin and the circumsporozoite protein of malaria parasites contain similar sequence motifs.

A;Reference number: S05478; MUID:88318954; PMID:3045564

A;Accession: S05478

A;Molecule type: mRNA

A;Residues: 1-437 <GOU>

A;Cross-references: UNIPROT:P11680; EMBL:X12905; NID:g53786; PIDN:CAA31389.1;
PID:g53787

C;Complex: a mixture of homodimers, homotrimers and homotetramers

C;Function:

A;Description: protects C3 convertase (C3bBb) from rapid inactivation

A;Pathway: complement alternate pathway

C;Superfamily: human properdin precursor; thrombospondin type 1 repeat homology

C;Keywords: complement alternate pathway; glycoprotein; homodimer; homotetramer;
homotrimer; plasma

F;45-97/Domain: thrombospondin type 1 repeat homology <THR1>

F;104-160/Domain: thrombospondin type 1 repeat homology <THR2>

F;161-224/Domain: thrombospondin type 1 repeat homology <THR3>

F;225-282/Domain: thrombospondin type 1 repeat homology <THR4>

F;283-345/Domain: thrombospondin type 1 repeat homology <THR5>

F;346-408/Domain: thrombospondin type 1 repeat homology <THR6>

F;52,55,108,111,114,165,168,229,232,290,293,350,353,356/Modified site: 2'-
mannosyl-tryptophan (Trp) #status predicted

F;366,396/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.8%; Score 229; DB 2; Length 437;
Best Local Similarity 40.4%; Pred. No. 9.9e-09;
Matches 46; Conservative 10; Mismatches 46; Indels 12; Gaps 4;

```
Qy      243 GGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACAT--LCPVDGS 300
      |||| | | || :| :| | | | |||| || | |: | || | || | :
```

Db 106 GGWSEWGPWGPCSVTCSKGTQIRQRVCDNPAPKCGG-HCPGEAQSQACDTQKTCPTHGA 164
Qy 301 WSPWSKWSACGLDC-----THWRSRECSDPAPRN--GGEECQGTDLDTNCT 345
|: | || | ||| ||| : |: | | : : |:
Db 165 WASWGPWSPRSGSCLGGAQEPKETRSRSCSAPAPSHQPPGKPCSGPAYEHKACS 218

Search completed: March 1, 2005, 09:07:16
Job time : 38.0558 secs

OM protein - protein search, using sw model

Run on: March 1, 2005, 08:46:18 ; Search time 35.0558 Seconds
(without alignments)
2464.715 Million cell updates/sec

Title: US-10-624-932-2
Perfect score: 4791
Sequence: 1 MAVRPGLWPALLGIVLAWL.....AVAGLGQPDAGLFTVSEAE 898

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	977	20.4	919	2	T32541	unc-5 protein - Ca
2	977	20.4	947	1	B44294	unc-5 protein, lon
3	298.5	6.2	1584	2	T00026	brain-specific ang
4	296.5	6.2	1172	1	TSHUP2	thrombospondin 2 p
5	293	6.1	1074	2	JC5928	semaphorin F precu
6	293	6.1	1172	2	A42587	thrombospondin 2 p
7	276	5.8	1444	2	T18856	angiogenesis inhib
8	275	5.7	984	2	T00326	hypothetical prote
9	275	5.7	1522	2	T00028	brain-specific ang
10	274.5	5.7	1572	2	T00027	brain-specific ang
11	270.5	5.6	1170	2	A40558	thrombospondin 1 p
12	268.5	5.6	1170	1	TSHUP1	thrombospondin 1 p
13	263	5.5	1178	1	A39804	thrombospondin pre

14	243	5.1	469	1	S29126	properdin precurs
15	229	4.8	437	2	S05478	properdin - mouse
16	226	4.7	254	2	T15952	hypothetical prote
17	221.5	4.6	1265	1	A37967	neural cell adhesi
18	215	4.5	788	2	T25061	hypothetical prote
19	215	4.5	1651	2	T14160	transmembrane rece
20	208	4.3	1612	2	T30805	dutt1 protein - mo
21	191.5	4.0	1344	2	T14316	rig-1 protein - mo
22	191	4.0	1863	2	S46217	protein-tyrosine-p
23	188.5	3.9	957	2	T15976	hypothetical prote
24	186	3.9	423	2	T29549	hypothetical prote
25	181.5	3.8	1273	2	T42405	sax-3 protein - Ca
26	181	3.8	1736	2	A47747	tight junction pro
27	178	3.7	1745	2	A46431	tight junction-ass
28	175	3.7	1907	2	S50893	protein-tyrosine-p
29	172	3.6	837	2	T00355	hypothetical prote
30	169.5	3.5	934	1	A34372	complement C6 prec
31	168.5	3.5	152	2	D89753	protein F11C7.2 [i
32	168	3.5	860	2	T16892	hypothetical prote
33	162.5	3.4	654	2	T29247	hypothetical prote
34	161	3.4	1501	2	I58148	protein-tyrosine-p
35	159.5	3.3	951	2	T00017	gene ADAMTS-1 prot
36	159	3.3	805	2	T34212	hypothetical prote
37	158.5	3.3	2165	2	T21371	hypothetical prote
38	157	3.3	550	2	T47158	hypothetical prote
39	156.5	3.3	807	2	A38152	F-spondin - rat
40	156	3.3	584	1	C8HUA	complement C8 alph
41	155	3.2	1499	2	I50212	protein-tyrosine-p
42	155	3.2	1898	2	S46216	leukocyte antigen-
43	150.5	3.1	1437	2	T31093	probable protein-t
44	149.5	3.1	712	2	A45638	immunodominant mic
45	148.5	3.1	206	2	A45517	coccidiosis-relate

ALIGNMENTS

RESULT 1

T32541

unc-5 protein - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T32541

R;Latreille, P.

submitted to the EMBL Data Library, December 1997

A;Description: The sequence of *C. elegans* cosmid B0273.

A;Reference number: Z21187

A;Accession: T32541

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-919 <LAT>

A;Cross-references: UNIPROT:O44171; EMBL:AF036698; PIDN:AAB88355.1;

GSPDB:GN00022; CESP:B0273.4a

A;Experimental source: strain Bristol N2; clone B0273

C;Genetics:

A;Gene: unc-5; CESP:B0273.4a

A;Map position: 4

A; Introns: 41/3; 108/1; 142/3; 201/1; 323/2; 553/1; 858/3
 C; Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology;
 thrombospondin type 1 repeat homology

Query Match 20.4%; Score 977; DB 2; Length 919;
 Best Local Similarity 28.7%; Pred. No. 2.3e-62;
 Matches 265; Conservative 168; Mismatches 379; Indels 110; Gaps 31;

Qy	49	EPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIER--STDGSSGLPTMEVRI	106
		: : :: : :: :: :: :: :: :: :: :: ::	
Db	9	QPKSGYVIRNKPLRLQCRANHATKIRYKCSSKWID--DSRIEKLIGTDSTSGVGYIDASV	66
Qy	107	NVSRQQVEKVFGLEEYWCQCVAWSSSG-----TTKSQKAYIRIARLRKNFEQEPLAKEVS	161
		:: :: :: :: :: :: :: :: :: :: :: :: ::	
Db	67	DISRIDVDTSGHVDAFQCQCYA---SGDDDQDVVASDVATVHLAYMRKHFLKSPVAQRVQ	123
Qy	162	LEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPNVDNVIYITREHSLVVRQARLADTANYTC	221
		: :: :: :: :: :: :: :: :: :: :: :: :: ::	
Db	124	EGTTLQLPCQAPESDPKAELTWYKDGVVVQP--DANVIRASDGLIMSAAARLSDSGNYTC	181
Qy	222	VAKNIVARRRSASAAVIVYVNGGWSTWTEW-SVCSASCG-----RGWQKR	265
		:: :: :: :: :: :: :: :: :: :: :: :: ::	
Db	182	EATNVANSRKTDPEVQIYVDGGWSEWSPWIGTCHVDCPLLQHAHRI RDPHDVLP HQRR	241
Qy	266	SRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP	325
		: : :: :: :: :: :: :: :: :: :: :: :: ::	
Db	242	TRTCNNPAPLNDGEYCKGEEEMTRSCVKPCKLDGGWSSWSWSDWSACSSSCHRYRTRACTVP	301
Qy	326	APRNGGEECQGTDL DTRNCTSDLCVHSASG--PEDVALYVGLIAVAVCLVLLLLVLILVY	383
		:: :: :: :: :: :: :: :: :: :: :: :: ::	
Db	302	PPMNGGQPCFGDDLMTQECPAQLCTADSSRIVISDTAVYGSVASIFIVASFILAILAMFC	361
Qy	384	CR-----KKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLLTI-----	422
		: :: :: :: :: :: :: :: :: :: :: :: :: ::	
Db	362	CKRGNSKKSPLKPQKMNSEKAGGIYYS---EPPGVRRLLEHQHGTLLGEKISSCSQYF	418
Qy	423	-QPDLSTTTT-----TYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHSSPT-SE	474
		:: :: :: :: :: :: :: :: :: :: :: :: ::	
Db	419	EPPPLPHSTTLRSKGSAFSGYSSTRNAGSRAALIQECSSSSSGSGGKRTMLRTSSSNCS D	478
Qy	475	AEEFVSRLSTQNYFRSLPRGTS-NMTYGTNFNLGGRMLIPNTGISLLIPDAIPRGKIYE	533
		: : : :: :: :: :: :: :: :: :: :: :: :: ::	
Db	479	DDNYATLYDYMEDKSVLGLDTSQNIVAAQIDSNGARLSLSKSGARLIVPELAVEGEKM--	536
Qy	534	IYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGV-----LLTRPVILAMDHCGEPS-P	585
		: : :: :: :: :: :: :: :: :: :: :: :: :: ::	
Db	537	LYLAVSDTLTDQPHLKPIESALSPVIVIGQCDVMSMSAHDNILRRPVVVSFRHCASTFPRD	596
Qy	586	SWSLRLKKQSCEGS-WEDVLHLGEEAPSHLYYCQLEASA-----CYVFTEQLGRFAL	636
		: :: :: :: :: :: :: :: :: :: :: :: :: ::	
Db	597	NWQFTL--YADEGSGWQKAVTIGEENLNTNMFVQFEQPGKKNDGFGWCHVMTYSLARLML	654
Qy	637	VGEAL--SVAAAKRLKLLLFPVACTSLE--YNIRVYCLHDTHDALKEVVQLEKQLGGQL	692
		:: :: :: :: :: :: :: :: :: :: :: :: ::	
Db	655	AGHPRRNSLSAAKRVHLAVFGPT EMSAYRRPFELRVYCVPETGAAMESVWKQED--GSRL	712
Qy	693	IQEPR--VLHFKDSYHNLRLSIHDV-PSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFT	749

```

      : |   :|: |   || : | || |   :| |   :   | ||:
Db      713 LCESNDFILNEKG---NLCICIEDVIPGFSCDGPEVVEISETQHRFV---AQNGLHCSLK 766

Qy      750 LERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAF 809
      : |   : :| :|   :   :   :   :   : | | |   :   |
Db      767 FRPKEINGSQFSTRVIVYQKASSTEPMVM--EVSNEPELYDATSEEREKGSVCV----EF 820

Qy      810 KIPFLIRQKIISSLDPPCRRGADWRTLQAKLHLDSHLSFFASKP--SPTAMILNLWEARH 867
      ::|| :: ::   || |   :||| ||:|||| | :| |||| |   |||:::|:||||
Db      821 RLPFGVKDELARLLDMPNESHSDWRGLAKKLHYDRYLQFFASFPDCSPTSLLLDLWEASS 880

Qy      868 FPNGN-LSQLAAAVAGLGQPPDA 888
      :   :   |   :   :|:||||
Db      881 SGSARAVPDLLQTLRVMGRPPDA 902

```

RESULT 2

B44294

unc-5 protein, long form - *Caenorhabditis elegans*

N;Contains: unc-5 protein, short form

C;Species: *Caenorhabditis elegans*

C;Date: 30-Apr-1993 #sequence_revision 28-Jul-1995 #text_change 09-Jul-2004

C;Accession: B44294; T32540; A44294

R;Leung-Hagesteijn, C.; Spence, A.M.; Stern, B.D.; Zhou, Y.; Su, M.W.;

Hedgecock, E.M.; Culotti, J.G.

Cell 71, 289-299, 1992

A;Title: UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 domains, guides cell and pioneer axon migrations in *C. elegans*.

A;Reference number: A44294; MUID:93046629; PMID:1384987

A;Contents: variety Bergerac

A;Accession: B44294

A;Molecule type: DNA

A;Residues: 1-947 <LEU>

A;Cross-references: UNIPROT:O44171; GB:S47168; NID:g258527; PIDN:AAB23867.1; PID:g258529

A;Note: sequence extracted from NCBI backbone (NCBIN:116668, NCBIN:116670, NCBIN:116672, NCBIN:116674, NCBIN:116676, NCBIN:116678, NCBIN:116680, NCBIN:116682, NCBIN:116685, NCBIP:118648)

A;Note: authors translated the codon CTA for residue 642 as Val; sequence shown follows the authors' translation

A;Note: mRNA lacking the first exon is equally prevalent

R;Latreille, P.

submitted to the EMBL Data Library, December 1997

A;Description: The sequence of *C. elegans* cosmid B0273.

A;Reference number: Z21187

A;Accession: T32540

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-947 <LAT>

A;Cross-references: EMBL:AF036698; PIDN:AAB88356.1; GSPDB:GN00022; CESP:B0273.4b

A;Experimental source: strain Bristol N2; clone B0273

C;Genetics:

A;Gene: unc-5

A;Map position: 4

A;Introns: 28/1; 69/3; 136/1; 170/3; 229/1; 351/2; 581/1; 886/3

C;Function:

A;Description: required for guidance of pioneering axons and cells migrating dorsally along the body wall; proposed to be a receptor on the surface of the motile cells

C;Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology; thrombospondin type 1 repeat homology

C;Keywords: alternative splicing; duplication; glycoprotein; receptor; transmembrane protein

F;30-947/Product: unc-5 protein, short form #status predicted <ALT>

F;46-116/Domain: immunoglobulin homology <IM1>

F;153-211/Domain: immunoglobulin homology <IM2>

F;229-300/Domain: thrombospondin type 1 repeat homology #status atypical <THR1>

F;301-354/Domain: thrombospondin type 1 repeat homology <THR2>

F;365-390/Domain: transmembrane #status predicted <TMM>

F;512-559/Domain: SH3 homology <SH3>

F;53-114,65-112,160-209/Disulfide bonds: #status predicted

F;206/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.4%; Score 977; DB 1; Length 947;

Best Local Similarity 28.7%; Pred. No. 2.4e-62;

Matches 265; Conservative 168; Mismatches 379; Indels 110; Gaps 31;

```
Qy      49 EPEDVYIVKKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIER--STDGSSGLPTMEVRI 106
      ||| |::||| | ||| ||| :|| :|| | ||| || :|| :
Db      37 QPKSGYVIRNKPLRLQCRANHATKIRYKCSSKWID--DSRIEKLIGTDSTSGVGYIDASV 94

Qy     107 NVSRQQVEKVFGLEEYWCQCVAWSSSG-----TTKSQKAYIRIARLRKNFEQEPLAKEVS 161
      :|| | : : ||| | || | : : | :||| : ||| |
Db      95 DISRIDVDTSGHVDAFQCQCYA---SGDDDQDVVASDVATVHLAYMRKHFLKSPVAQRVQ 151

Qy     162 LEQGIVLPCRPPGEGIPPAEVEWLRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTC 221
      : ||| : || | ||| : : : || | || | : ||| : ||| : |||
Db     152 EGTTLQLPCQAPESDPKAELTWYKDGVVVQP--DANVIRASDGLIMSAAARLSDSGNYTC 209

Qy     222 VAKNIVARRRSASAAVIVYVNGGWSTWTEW-SVCSASCG-----RGWQKR 265
      | : | : | : ||| || : | | | | | | | | | | |
Db     210 EATNVANSRKTDPEVQIYVDGGWSEWSPWIGTCHVDCPLLRQHAHRIRDPHDVLPHQRR 269

Qy     266 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP 325
      :|| | |||| | :|| : : | | :|| || || || || | :|| | : |
Db     270 TRTCNNPAPLNDGEYCKGEEEMTRSCVKPCKLDGGWSSWSWSACSSSCHRYRTRACTVP 329

Qy     326 APRNGGEECQGTDLDTNRCTSDLCVHSASG--PEDVALYVGLIAVAVCLVLLLLVLILVY 383
      | ||| : | | ||| : | : || : | | | : : : : : : : | : | :
Db     330 PPMNGGQPCFGDDLMTQECPAQLCTADSSRIVISDTAVYGSVASIFIVASFILAILAMFC 389

Qy     384 CR-----KKEGLSDVDADSSILTSGFQPVSIKPSKADNPHLLTI----- 422
      | : | : : ||| : : : | : : : : : : : : : : :
Db     390 CKRGNSKSKPLKPQKMNSEKAGGIYYS---EPPGVRRLLEHQHGTLLGEKISSCSQYF 446

Qy     423 -QPDLSSTTTT-----TYQGSILCPQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPT-SE 474
      | | :|| : | | | | | | | | | | | | | | | |
Db     447 EPPPLPHSTTLRSGKSAFSGYSSTRNAGSRAALIQECSSSSSGSGGKRTMLRTSSSNCS 506

Qy     475 AEEFVSRLSTQNYFRSLPRGTS-NMTYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYE 533
      : : : | || : : | || : : | :|| | : | :
Db     507 DDNYATLYDYMEDKSVLGLDTSQNIVAAQIDSNGARLSLSKSGARLIVPELAVEGEKM-- 564
```

Qy 534 IYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGV-----LLTRPVILAMDHCGEPSP-D 585
 :|| : : | ::|||: | | :| |||:: || | |
 Db 565 LYLAVSDTLTDQPHLKPIESALSPVIVIGQCDVSMSAHDNILRRPVVVSFRHCASTFPRD 624

Qy 586 SWSLRLKKQSCEGS-WEDVLHLGEEAPSHLYYCQLEASA-----CYVFTEQLGRFAL 636
 :| | : ||| |: : :||| : : | | :| | | | |
 Db 625 NWQFTL--YADEGSGWQKAVTIGEENLNTNMFVQFEQPGKKNDFGWCHVMTYSLARLML 682

Qy 637 VGEAL--SVAAAKRLKLLLFPVACTSLE--YNIRVYCLHDTHDALKEVVQLEKQLGGQL 692
 | :|||: | :| | : : :|||: :| |:: | : | :|
 Db 683 AGHPRRNSLSAAKRVHLAVFGPTMSAYRRPFELRVYCVPETGAAMESVWKQED--GSRL 740

Qy 693 IQEPR--VLHFKDSYHNLRSLSIHDV-PSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFT 749
 : | :|: | || : | || | :| | : | |||:
 Db 741 LCESNDFILNEKG---NLCICIEDVIPGFSCDGPVEVEISETQHRFV---AQNLHCSLK 794

Qy 750 LERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAF 809
 : | : :: |:| : : : : : : | | | : |
 Db 795 FRPKEINGSQFSTRVIVYQKASSTEPMVM--EVSNEPELYDATSEEREKGSVCV----EF 848

Qy 810 KIPFLIRQKIISLDPPCRRGADWRTLAQKLHLDShLSFFASKP--SPTAMILNLWEARH 867
 ::|| :: :: || | :|| ||:|| | :| ||| | |||::|:||||
 Db 849 RLPFGVKDELARLLDMPNESHSDWRGLAKKLHYDRYLQFFASFPDCSPTSLLLDLWEASS 908

Qy 868 FPNGN-LSQLAAAVAGLQGPDA 888
 : : | : :|||
 Db 909 SGSARAVPDLQLTLRVMGRPDA 930

RESULT 3

T00026

brain-specific angiogenesis inhibitor 1 - human

N;Alternate names: BAI1 protein

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004

C;Accession: T00026

R;Nishimori, H.; Shiratsuchi, T.; Urano, T.; Kimura, Y.; Kiyono, K.; Tatsumi, K.; Yoshida, S.; Ono, M.; Kuwano, M.; Nakamura, Y.

submitted to the EMBL Data Library, June 1997

A;Reference number: Z14064

A;Accession: T00026

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1584 <NIS>

A;Cross-references: UNIPROT:O14514; EMBL:AB005297; NID:d1175078; PID:d1024528

A;Experimental source: brain

C;Genetics:

A;Gene: GDB:BAI1

A;Cross-references: GDB:9838088; OMIM:602682

A;Map position: 8q24-8q24

F;408-462/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 6.2%; Score 298.5; DB 2; Length 1584;

Best Local Similarity 33.5%; Pred. No. 5.2e-13;

Matches 78; Conservative 35; Mismatches 91; Indels 29; Gaps 11;

Qy 124 CQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEW 183

```

      | | :| | |: :| | : :| | | | | | |
Db      309 CNREACGPAGRTSSRSQSLRSTDARR---REELGDEL---QQFGFPA-PQTGDPAAE-EW 360

Qy      184 LRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNG 243
      : | | : || | : ::| : :| : : : | :|
Db      361 --SPWSVCSSTCGEGWQTR-----TRFCVSSSYSTQCSGPLREQRLCNSAVCPVHG 410

Qy      244 GWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPLNGGAFCEGQNVQKTAC-ATLCP---VDG 299
      | |: ||: ||: ||: ||: ||: || | | | | | | | | | | |
Db      411 AWDEWSPWSLCSSTCGRGRDRTRTCR--PPQFGGNPCEGPEKQTKFCNIALCPGRAVDG 468

Qy      300 SWSPWSKWSACGLDCT---HWSRECSDPAPRNGGEECQGTDLDTNRNCTSDLC 349
      :|: || |||| | : | :||: | : || |||| : :||: | |
Db      469 NWNEWSSWSACSASCSQGRQRTRECNGPS--YGGAECQGHVETRDCLQQC 519

```

RESULT 4

TSHUP2

thrombospondin 2 precursor - human

C;Species: Homo sapiens (man)

C;Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C;Accession: A47379; A42173

R;LaBell, T.L.; Byers, P.H.

Genomics 17, 225-229, 1993

A;Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: potential regulatory role for the 3' untranslated region.

A;Reference number: A47379; MUID:94010892; PMID:8406456

A;Accession: A47379

A;Molecule type: mRNA

A;Residues: 1-1172 <LAB>

A;Cross-references: UNIPROT:P35442; GB:L12350; NID:g307505; PIDN:AAA03703.1; PID:g307506

R;LaBell, T.L.; Milewicz, D.J.; Disteché, C.M.; Byers, P.H.

Genomics 12, 421-429, 1992

A;Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression of a second member of the thrombospondin gene family in humans.

A;Reference number: A42173; MUID:92217961; PMID:1559694

A;Accession: A42173

A;Molecule type: mRNA

A;Residues: 560-1172 <LA2>

A;Cross-references: GB:M81339

A;Experimental source: fibroblast

A;Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)

C;Genetics:

A;Gene: GDB:THBS2; TSP2

A;Cross-references: GDB:128789; OMIM:188061

A;Map position: 6q27-6q27

C;Complex: homotrimer, disulfide linked

C;Function:

A;Description: participates in cell migration and adhesion, and in platelet aggregation

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-1172/Product: thrombospondin 2 #status predicted <MAT>

F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>
 F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
 F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
 F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
 F;553-588/Domain: EGF homology <EGF1>
 F;652-691/Domain: EGF homology <EGF>
 F;928-930/Region: cell attachment (R-G-D) motif
 F;151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent)
 #status predicted
 F;167-226/Disulfide bonds: #status predicted
 F;266,270/Disulfide bonds: interchain #status predicted
 F;612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 6.2%; Score 296.5; DB 1; Length 1172;
 Best Local Similarity 30.5%; Pred. No. 4.9e-13;
 Matches 78; Conservative 28; Mismatches 105; Indels 45; Gaps 9;

Qy	209	RQARLADTANYTCVAKNIVARRRSASAA-VIVYVNGGWSTWTEWSVCSASCGRGWQKRSR	267
		: : : : : : : :	
Db	403	QRGRSCDVTSNTCLGPSIQTRACSLSKCDTRIRQDGGWSHWSPWSSCSVTCGVGNITRIR	462
Qy	268	SCTNPAPLNGGAFCEGQNVQKTAC-ATLCPVDGWSWSPWSKWSACGLDCT---HWRSRECS	323
		: : : : : :	
Db	463	LCNSPVPQMGGKNCKGSGRETAKCQGAPCPIDGRWSPWSPWSACTVTCAGGIRERTRVCN	522
Qy	324	DPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVY	383
		: : :	
Db	523	SPEPQYGGKACVGDVQERQMCNKRSC-----PVDGCLSNPCFPGAQC-----	564
Qy	384	CRKKEGLDSDVADSSILTSGFQPVSI--KPSKADNPHLLTIQPDLSSTTTT-----TYQ	434
		: : : : : :	
Db	565	-----SSFPGDS-WSCGFPCVPGFLNGTHCEDLDECALVPDICFSTSKVPRCVNTQP	615
Qy	435	GSLC----PRQDGPSP	446
Db	616	GFHCLPCPPRYRGNQP	631

RESULT 5

JC5928

semaphorin F precursor - human

C;Species: Homo sapiens (man)

C;Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

C;Accession: JC5928

R;Simmons, A.D.; Pueschel, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.
 Biochem. Biophys. Res. Commun. 242, 685-691, 1998

A;Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-chat candidate interval.

A;Reference number: JC5928; MUID:98125554; PMID:9464278

A;Accession: JC5928

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-1074 <SIM>

A;Cross-references: UNIPROT:Q13591; GB:U52840; NID:g2772583; PIDN:AAC09473.1;
 PID:g2772584

A;Experimental source: brain

F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
 F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
 F;553-588/Domain: EGF homology <EGF1>
 F;652-691/Domain: EGF homology <EGF>

Query Match 6.1%; Score 293; DB 2; Length 1172;
 Best Local Similarity 38.0%; Pred. No. 8.7e-13;
 Matches 60; Conservative 22; Mismatches 66; Indels 10; Gaps 5;

```

Qy      209 RQARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGGWSTWTEWSVCASCGRGWQKRSR 267
          :: | | : ||: :| | | : |||| | : || || :|| | | |
Db      403 QRGRSCDVTSNTCLGPSIQTRTCSLGKCDTRIRQNGGWSHWSWPSSCSVTGCVGNVTRIR 462

Qy      268 SCTNPAPLNGGAFCEGQNVQKTAC-ATLCPVDGSWSPWSKWSACGLDCT---HWRSRECS 323
          | :| | || | :| : | ||: || |||| |||| : | || | :
Db      463 LCNSPVPQMGGKNCKGSGRETQPCQDPCPIDGRWSPWSPWSACTVTCAGGIRERSRVCN 522

Qy      324 DPAPRNGGEECQG--TD---LDTRNCTSDLCVHSASGP 356
          | : ||: || | | : : | : | | : : |
Db      523 SPEPQYGGKDCVGDVTEHQMCNKRSCPIDGCLSNPCFP 560

```

RESULT 7

T18856

angiogenesis inhibitor homolog - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T18856; T24653

R;McMurray, A.

submitted to the EMBL Data Library, July 1995

A;Reference number: Z19031

A;Accession: T18856

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1444 <WIL>

A;Cross-references: UNIPROT:Q8MYA8; EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1

A;Experimental source: clone C02B4

R;McMurray, A.

submitted to the EMBL Data Library, July 1995

A;Reference number: Z19917

A;Accession: T24653

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1444 <WI2>

A;Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1

A;Experimental source: clone T07C5

C;Genetics:

A;Gene: CESP:C02B4.1

A;Map position: X

A;Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3;

509/3; 566/2; 625/1; 696/2; 786/3; 812/2; 878/3; 971/1; 1007/3; 1067/1; 1099/3;

1180/3; 1273/2; 1305/1; 1363/1; 1388/2

Query Match 5.8%; Score 276; DB 2; Length 1444;
 Best Local Similarity 27.2%; Pred. No. 2e-11;
 Matches 73; Conservative 28; Mismatches 97; Indels 70; Gaps 12;

Qy 123 WCQCVAWSSSGTTKSQKAYIRIARLRKNFEQ-----EPLAKEVSLEQGIVLPCRPEGI 176
 | : :||: : : | : | : ||| :|| |
 Db 1134 WSESSWSAC-----SCFSLTSTRRRFCQVVDPTVQGFCAGAI EQ---IPCAPGSCS 1183

Qy 177 PPAE-----VEW-----LRNEDLVDP SLDPNVYITREHSLVVRQARLADTAN 218
 | | || :|| :| : | | :
 Db 1184 PSAGGWSLWSEWSSCSKDCGDTGHQIRNRM CSEP-----IPSNRGAYCSG 1228

Qy 219 YT-----CVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR SR SCTNPA 273
 |: || |: : : |:||: || || |: | | : |:| | |
 Db 1229 YSFDQRPCVMDNVCSDEK-----VDGGWTDWTAWSECTDYCRNGHRSRTRFCANPK 1279

Qy 274 PLNGGAFCEGQNVQKTAC--ATLCPV-DGSWSPWSKWSACGLDC---THWRSRECS DPAP 327
 | ||| | | : : | : || || || |: | | | ||| | |
 Db 1280 PSQGGAQCTGSDFELNPCFDPARCHLRDGGWSTWSDWT PCSASC GFGVQTRDRSCSSPEP 1339

Qy 328 RNGGEECQGTDL DTRNCTSDLCVHSASG 355
 : ||: | | | | | : |
 Db 1340 K-GGQSCSGLAHQTS LCDLPACDHESDG 1366

RESULT 8

T00326

hypothetical protein KIAA0550 - human

C;Species: Homo sapiens (man)

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 15-Mar-2004

C;Accession: T00326

R;Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.

DNA Res. 5, 31-39, 1998

A;Title: Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.

A;Reference number: Z14086; MUID:98290545; PMID:9628581

A;Accession: T00326

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-984 <NAG>

A;Cross-references: EMBL:AB011122; NID:g3043623; PIDN:BAA25476.1; PID:g3043624

A;Experimental source: brain

C;Genetics:

A;Note: KIAA0550

F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match . . . 5.7%; Score 275; DB 2; Length 984; . . .

Best Local Similarity 39.0%; Pred. No. 1.4e-11;

Matches 57; Conservative 20; Mismatches 53; Indels 16; Gaps 6;

Qy 220 TCVA-----KNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR SR SCTNPA 273
 |||: : | : : |:| | |: ||:| | :||| : |:|||
 Db 317 TCVSPYGT HCSGPLRESRVCNNTALCPVHGVWEEWSPWSLCSFTCGRGQRT RTR SCT--P 374

Qy 274 PLNGGAFCEGQNVQKTAC--ATLCPVDGSWSPWSKWSACGLDC---THWRSRECS DPAPRN 329
 | || ||| | ||||| | || || | : | | |||:| : | :
 Db 375 PQYGGRPCEGPETHHKPCNIALCPVDGQWQEWSWSQCSVTCSNGTQQRSRQCT--AAA H 432

Qy 330 GGEQCQGTDLDTNRCTSDLCVHSASG 355
 || ||:| :| | : | :|:|
 Db 433 GGSECRGPWAESRECYNPEC--TANG 456

RESULT 9

T00028

brain-specific angiogenesis inhibitor 3 - human

N;Alternate names: BAI3 protein

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004

C;Accession: T00028

R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.

Cytogenet. Cell Genet. 79, 103-108, 1997

A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-specific angiogenesis inhibitor 1 (BAI 1).

A;Reference number: Z14066; MUID:98194217; PMID:9533023

A;Accession: T00028

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1522 <SHI>

A;Cross-references: UNIPROT:O60242; EMBL:AB005299; NID:g3021700;

PIDN:BAA25363.1; PID:g3021701

A;Experimental source: brain

C;Genetics:

A;Gene: GDB:BAI3

A;Cross-references: GDB:9838090; OMIM:602684

A;Map position: 6q12-6q12

F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 5.7%; Score 275; DB 2; Length 1522;

Best Local Similarity 39.0%; Pred. No. 2.5e-11;

Matches 57; Conservative 20; Mismatches 53; Indels 16; Gaps 6;

Qy 220 TCVA-----KNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCNPA 273

|||: : | : : | :| | | : ||:| : ||| : | : |||

Db 317 TCVSPYGTGTHCSGPLRESRVCNNTALCPVHGVEEWSPWSLCSFTCGRQRTTRTRST--P 374

Qy 274 PLNGGAFCEGQNVQKTAC-ATLCPVDGWSWSPWSKWSACGLDC---THWRSRECSDPAPRN 329

| || ||| | ||||| | || || : | | |||:| : | :

Db 375 PQYGGRPCEGPEETHHKPCNIALCPVDGQWQEWSSWSQCSVTCSNGTQQRSRQCT--AAAH 432

Qy 330 GGEQCQGTDLDTNRCTSDLCVHSASG 355

|| ||:| :| | : | :|:|

Db 433 GGSECRGPWAESRECYNPEC--TANG 456

RESULT 10

T00027

brain-specific angiogenesis inhibitor 2 - human

N;Alternate names: BAI2 protein

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004

C;Accession: T00027

R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.

Cytogenet. Cell Genet. 79, 103-108, 1997

A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-specific angiogenesis inhibitor 1 (BAI 1).
 A;Reference number: Z14066; MUID:98194217; PMID:9533023
 A;Accession: T00027
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-1572 <SHI>
 A;Cross-references: UNIPROT:O60241; EMBL:AB005298; NID:g3021698;
 PIDN:BAA25362.1; PID:g3021699
 A;Experimental source: brain
 C;Genetics:
 A;Gene: GDB:BAI2
 A;Cross-references: GDB:9838089; OMIM:602683
 A;Map position: 1p35-1p35

Query Match 5.7%; Score 274.5; DB 2; Length 1572;
 Best Local Similarity 19.2%; Pred. No. 2.8e-11;
 Matches 176; Conservative 108; Mismatches 307; Indels 327; Gaps 38;

Qy	173	PEGIPPAEVEWLRNEDLVDPSPDPNVY-----ITREHSLVVRQARL	213
		: : : : : :	
Db	271	PEEEPVKVTQWPRAD-----EPGLYMAQTGDPAEEWSPWSVCSLTGQGLQVR-TRS	323
Qy	214	ADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCNPA	273
		: : : : : : : : :	
Db	324	CVSSPYGTLCSGPLRETRPCNNSATCPVHGVWEEWGSWSLCSRSCGRGSRSRMRTCV--P	381
Qy	274	PLNGGAFCEGQNVQKTACA-TLCPVDGSWSPWSKWSACGLDC---THWRSRECSDPAPR-	328
		: : : :	
Db	382	PQHGGKACEGPELQTKLCSMAACPVEGQWLEWGPWGPCSTSCANGTQQRSRKCSVAGPAW	441
Qy	329	-----NGGEECQ	335
Db	442	ATCTGALTDTRECSNLECPATDSKWGPWNAWSLCSKTCDTGWQRRFRMCQATGTQGYPC	501
Qy	336	GTDLDTNRCTSDLC--VHSASGPEDVAL-----	361
		: : :	
Db	502	GTGEEVKPCSEKRCPAFHEMCRDEYVMLMTWKKAAGEIIYNKCPPNASGSASRRCLLSA	561
Qy	362	----YVGLIAVAVCL---VLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKA	414
		: : : : : : : : : : : : : :	
Db	562	QGVAYWGLPSFARCISHEYRYLYLSLREHLAKGQRMLAGEGMSQVVRSLQELLARRTYY	620
Qy	415	DNPHLLTIQPDLSSTTTTQYQGSCLPRQDGPSPKFQLT-----NGHLLSPLGG	461
		: : : : : : : : :	
Db	621	SGDLLFSVDILRNVTDTFKRATYVPSADDVQRFQVVSFMVDAENKEKWDDAQVSP--G	678
Qy	462	GRHTLHHSSPTSEAEFV-----SRLSTQNYFRSLPRG-----TSNMTYGTFN	504
		: : : : : : : : : :	
Db	679	SVHLLR-----VVEDFIHLVGDAKAFQSSLIIVTDNLVISIQREPVS AVSSDITFPMRG	732
Qy	505	FLG-----GRLMIPNTGISLLIP-----PDAIPRGK-----	530
		: : : :	
Db	733	RRGMKDWVRHSEDRLFLPKEVLSLSSPGKPATSGAAGSPGRGRGP GTVPPGPGHSHQRL	792
Qy	531	-----IYE-IYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVIL	574
		: : : : :	

Db 793 PADPDESSYFVIGAVLYRTLGLILPPP---RPPLAVTSRVMT--VTVRPPTQPPAEPLIT 847
 Qy 575 A-----MDHCGEPPSPDSWSLRLKKQSCGWSWEDVLHLGEEAPSHLYYCQ-LEASACYV-- 626
 :: :| || : : | | : || || | :
 Db 848 VELSYIINGTTDPHCASWDYS-RADASSGDWD-----TENCQTLETQAAHTRC 894
 Qy 627 FTEQLGRFALVGE-----ALSVAANKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEV 681
 : | ||: : | : : | : : | : : | :
 Db 895 QCQHLSTFAVLAQPPKDLTLELAGSPSVPLVIGCAVSCMALLTLLAIYA-----AFWRF 948
 Qy 682 VQLEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDPSSLWKSLLVSYQEIPFYHIWNGTQ 741
 :: | : : | || : | : : || : : |
 Db 949 IKSERSI-----ILLNFCLSI--LASNI---LILVGQSRVLSKGVCTMTA 988
 Qy 742 RYLHCTFTLERVSPSTSDLACKLWV-----WQVEGDG 773
 : || | : || : || :
 Db 989 AFLHFFF-----LSSFCWVLTEAWQSYLAVIGRMRTLVRKRFLCLGWGLPALV 1037
 Qy 774 QSFSINFNITKDTFAELLALESEAG-VPALVGPSA-----FKIPFLIRQKI-----IS 821
 : | : | || : | | : | || : | : : | : ||
 Db 1038 VAVSVGFTRTKGYGTSSYCWSLEGGLLYAFVGPAAVIVLVNMLIGIIVFNKLMARDGIS 1097
 Qy 822 SLDPPCRRGAD---WRTL 836
 | | : : | : |
 Db 1098 DKSKKQRAGSERCPWASL 1115

RESULT 11

A40558

thrombospondin 1 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004

C;Accession: A40558; A37905; B42587; S68787

R;Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.

Genomics 11, 587-600, 1991

A;Title: Characterization of the murine thrombospondin gene.

A;Reference number: A40558; MUID:92128941; PMID:1774063

A;Accession: A40558

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1170 <LAW>

A;Cross-references: UNIPROT:P35441; GB:M62449; GB:M62450; GB:M62451; GB:M62452;
 GB:M62453; GB:M62454; GB:M62455; GB:M62456; GB:M62457; GB:M62458; GB:M62459;
 GB:M62460; GB:M62461; GB:M62462; GB:M62463; GB:M62464; GB:M62465; GB:M62466;
 GB:M62467; GB:M62468; GB:M62469; GB:M62470; NID:g511867; PIDN:AAA50611.1;
 PID:g511869

R;Bornstein, P.; Alfi, D.; Devarayalu, S.; Framson, P.; Li, P.

J. Biol. Chem. 265, 16691-16698, 1990

A;Title: Characterization of the mouse thrombospondin gene and evaluation of the role of the first intron in human gene expression.

A;Reference number: A37905; MUID:90375546; PMID:2398070

A;Accession: A37905

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-490 <BOR>

A;Cross-references: GB:J05605; GB:J05606; NID:g201991; PIDN:AAA40431.1; PID:g554390
R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M. J. Biol. Chem. 267, 3274-3281, 1992
A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell growth and development.
A;Reference number: A42587; MUID:92147683; PMID:1371115
A;Accession: B42587
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1152,'P',1154-1170 <LAH>
A;Cross-references: GB:M87276
A;Note: sequence extracted from NCBI backbone (NCBIP:81501)
R;Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F. FEBS Lett. 387, 36-41, 1996
A;Title: Expression and initial characterization of recombinant mouse thrombospondin 1 and thrombospondin 3.
A;Reference number: S68787; MUID:96234006; PMID:8654563
A;Accession: S68787
A;Molecule type: protein
A;Residues: 19-26,'X',28-37 <CHE>
C;Complex: homotrimer, disulfide linked
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology
C;Keywords: calcium binding; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF>
F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.6%; Score 270.5; DB 2; Length 1170; .
Best Local Similarity 32.2%; Pred. No. 3.7e-11;
Matches 57; Conservative 24; Mismatches 71; Indels 25; Gaps 5;

```

Qy      207 VVRQARLADTANYTCVAKNIVAR-----RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
          : :: | |: | | :: | :| :||| |: || || :|
Db      399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-----DGGWSHWPWSSCSVTC 451

Qy      259 GRGWQKRSRSCNPNAPLNGGAFCEGQNVQKTAC-ATLCPVDGSWSPWSKWSACGLDC--- 314
          | | | | | :|:| | |||: : || ||::| | || | | : |
Db      452 GDGVITRIRLCNSPSPQMNGKPCGEARETKACKKDACPINGGWGPWSPWDICSVTCGGG 511

Qy      315 THWRSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEdVALYVGLIAVAVC 371
          ||| |:| |: ||::| | : : | | | | | | | |
Db      512 VQRRSRLCNPPTPQFGGKDCVGDVTENQVCNKQDC-----PIDGCLSNPCFAGAKC 562

```

RESULT 12

TSHUP1

thrombospondin 1 precursor - human

C;Species: Homo sapiens (man)

C;Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C;Accession: A26155; A34274; A30140; A25812; A05172; A42927

R;Lawler, J.; Hynes, R.O.
 J. Cell Biol. 103, 1635-1648, 1986
 A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.
 A;Reference number: A26155; MUID:87057617; PMID:2430973
 A;Accession: A26155
 A;Molecule type: mRNA
 A;Residues: 1-1170 <LAW>
 A;Cross-references: UNIPROT:P07996; GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing
 R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.
 J. Biol. Chem. 264, 11222-11227, 1989
 A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.
 A;Reference number: A34274; MUID:89291870; PMID:2544587
 A;Accession: A34274
 A;Molecule type: DNA
 A;Residues: 1-166 <LAH>
 A;Cross-references: GB:J04835
 R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Frazier, W.A.
 J. Cell Biol. 108, 729-736, 1989
 A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.
 A;Reference number: A30140; MUID:89139590; PMID:2918029
 A;Accession: A30140
 A;Molecule type: mRNA
 A;Residues: 1-83, 'A', 85-522, 'A', 524-1170 <HEN>
 A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing
 R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.
 Biochemistry 25, 8418-8425, 1986
 A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.
 A;Reference number: A25812; MUID:87157592; PMID:3030396
 A;Accession: A25812
 A;Molecule type: mRNA
 A;Residues: 1-83, 'A', 85-397 <KOB>
 A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354
 R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.
 Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
 A;Reference number: A05172; MUID:86287276; PMID:3461443
 A;Accession: A05172
 A;Molecule type: mRNA
 A;Residues: 1-83, 'A', 85-374, 'RC' <DIX>
 A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing
 R;Sun, X.; Skorstengaard, K.; Mosher, D.F.
 J. Cell Biol. 118, 693-701, 1992
 A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.
 A;Reference number: A42927; MUID:92348511; PMID:1379247
 A;Accession: A42927

A;Molecule type: protein
 A;Residues: 987-1003 <SUN>
 A;Note: Cys-992 is shown to have a free sulfhydryl
 C;Genetics:
 A;Gene: GDB:THBS1; TSP1; TSP
 A;Cross-references: GDB:120438; OMIM:188060
 A;Map position: 15q15-15q15
 A;Introns: 23/1
 A;Note: the list of introns may be incomplete
 C;Complex: homotrimer, disulfide linked
 C;Function:
 A;Description: participates in cell migration and adhesion, and in platelet aggregation
 C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology
 C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
 F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
 F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
 F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
 F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
 F;551-586/Domain: EGF homology <EGF1>
 F;650-689/Domain: EGF homology <EGF2>
 F;926-928/Region: cell attachment (R-G-D) motif
 F;171-232/Disulfide bonds: #status predicted
 F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;270,274/Disulfide bonds: interchain #status predicted
 F;610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
 F;1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 5.6%; Score 268.5; DB 1; Length 1170;
 Best Local Similarity 32.9%; Pred. No. 5.1e-11;
 Matches 51; Conservative 24; Mismatches 61; Indels 19; Gaps 4;

```

Qy      207 VVRQARLADTANYTCVAKNIVAR-----RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
      : :: | | : | | :: | : | | | | | : |
Db      399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQCEDKRFKQ-----DGGWSHWSWPSSCSVTC 451

Qy      259 GRGWQKRSRSCNPNAPLNGGAFCEGQNVQKTAC-ATLCPVDGSWSPWSKWSACGLDC--- 314
      | | | | | : : | | | : | | : | | | | | : |
Db      452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKKDACPINGGWGPWSPWDICSVTCGGG 511

Qy      315 THWRSRECSDPAPRNGGEECQGTDLDTNRCTSDLC 349
      | | | | : : | | : | : | : | : | : |
Db      512 VQKRSRLCNNPTPQFGGKDCVGDVTENQICNKQDC 546
  
```

RESULT 13

A39804

thrombospondin precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C;Accession: A39804

R;Lawler, J.; Duquette, M.; Ferro, P.

J. Biol. Chem. 266, 8039-8043, 1991

A;Title: Cloning and sequencing of chicken thrombospondin.
 A;Reference number: A39804; MUID:91217026; PMID:2022631
 A;Accession: A39804
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1178 <LAW>
 A;Cross-references: UNIPROT:P35440; GB:M60853; NID:g212763; PIDN:AAA51437.1; PID:g212764
 C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology
 F;325-383/Domain: von Willebrand factor type C repeat homology <VWC>
 F;386-437/Domain: thrombospondin type 1 repeat homology <THR1>
 F;442-498/Domain: thrombospondin type 1 repeat homology <THR2>
 F;499-555/Domain: thrombospondin type 1 repeat homology <THR3>
 F;658-697/Domain: EGF homology <EGF>

Query Match 5.5%; Score 263; DB 1; Length 1178;
 Best Local Similarity 36.2%; Pred. No. 1.3e-10;
 Matches 58; Conservative 16; Mismatches 70; Indels 16; Gaps 5;

Qy	210	QARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRS	268
		: : :	
Db	410	RGRSCDVTRSACTGPHIQTRMCSFKKCDHRIRQDGGWSHWSPWSSCSVTCGVGNITRIRL	469
Qy	269	CTNPAPLNGGAFCEGQNVQKTACATL-CPVDGSWSPWSKWSACGLDC---THWRSRECS	324
		: : :	
Db	470	CNSPIPQMGGKNCVGNGRETEKCEKAPCPVNGQWGPWSPWSACTVTCGGGIRERSRLCNS	529
Qy	325	PAPRNGGEECQGTDLDT-----RNCTSDLCVHSASGP	356
		: : :	
Db	530	PEPQYGGKPCVG---DTKQHDMCNKRDCPIDGCLSNPCFP	566

RESULT 14
 S29126
 properdin precursor [validated] - human
 N;Alternate names: factor P
 C;Species: Homo sapiens (man)
 C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
 C;Accession: S29126; S16150; A05319; T45112; T45113
 R;Nolan, K.F.; Kaluz, S.; Higgins, J.M.G.; Goundis, D.; Reid, K.B.M.
 Biochem. J. 287, 291-297, 1992
 A;Title: Characterization of the human properdin gene.
 A;Reference number: S29126; MUID:93038568; PMID:1417780
 A;Accession: S29126
 A;Molecule type: DNA
 A;Residues: 1-469 <NOL1>
 A;Cross-references: UNIPROT:P27918; EMBL:X70872; NID:g35679; PIDN:CAA50220.1; PID:g35680
 R;Nolan, K.F.; Schwaeble, W.; Kaluz, S.; Dierich, M.P.; Reid, K.B.M.
 Eur. J. Immunol. 21, 771-776, 1991
 A;Title: Molecular cloning of the cDNA coding for properdin, a positive regulator of the alternative pathway of human complement.
 A;Reference number: S16150; MUID:91184288; PMID:2009915
 A;Accession: S16150
 A;Molecule type: mRNA
 A;Residues: 1-456, 'R', 458-469 <NOL2>

A;Cross-references: EMBL:X57748

R;Reid, K.B.M.; Gagnon, J.

Mol. Immunol. 18, 949-959, 1981

A;Reference number: A05319; MUID:82195224; PMID:7341961

A;Accession: A05319

A;Molecule type: protein

A;Residues: 28-53,'Q',55-59,'G',61,'I',63;137-138,'P',140-141,'P',143-144,'X',146-148,'Y',150,'S',152,'Y',154-156,'XSXGXA';162-163,'E',165-172,'X',174-176,'X',178,'V',180;223-228,'X',230-232,'GX',235-238,'GH',241-245;248-251,'X',253-257,'P',259,'G',261,'XPP',265-266,'X',268-269;280-285,'X',287-290,'X',292,'H',294-300,'SXXX',305-307,'X',309-315,'K',317;333-341,343-357,'X',359-362,'EXE';393-404,'QK',407;421-427,'R',429-443,'TKV',447-448,'XX',451,'RX',454-455 <REI>

R;Westberg, J.; Nordin-Fredrikson, G.; Truedsson, L.; Sjöholm, A.G.; Uhlen, M. submitted to the EMBL Data Library, May 1997

A;Reference number: Z22914

A;Accession: T45112

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-54,'X',56-73,'X',75-99,'W',101-469 <WES1>

A;Cross-references: EMBL:AF005665; PIDN:AAB63280.1

A;Experimental source: genomic DNA from individual with properdin deficiency type II

A;Accession: T45113

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-60,'X',62-413,'D',415-452,'XX',455-469 <WE2>

A;Cross-references: EMBL:AF005666; PIDN:AAC51626.1

A;Experimental source: genomic DNA from individual with properdin deficiency type III

R;Hartmann, S.; Hofsteenge, J.

J. Biol. Chem. 275, 28569-28574, 2000

A;Title: Properdin, the positive regulator of complement, is highly C-mannosylated.

A;Reference number: A59360; MUID:20435812; PMID:10878002

A;Contents: annotation

A;Note: identification and location of C-mannosylation sites by mass-spectroscopy

C;Genetics:

A;Gene: GDB:PFC

A;Cross-references: GDB:120275; OMIM:312060

A;Map position: Xp11.3-Xp11.23

A;Introns: 26/1; 76/2; 135/1; 192/1; 256/1; 314/1; 378/1; 415/2

C;Complex: a mixture of homodimers, homotrimers and homotetramers

C;Function:

A;Description: protects C3 convertase (C3bBb) from rapid inactivation

A;Pathway: complement alternate pathway

C;Superfamily: human properdin precursor; thrombospondin type 1 repeat homology

C;Keywords: complement alternate pathway; glycoprotein; homodimer; homotetramer; homotrimer; plasma

F;1-27/Domain: signal sequence #status predicted <SIG>

F;28-469/Product: properdin #status experimental <MAT>

F;76-128/Domain: thrombospondin type 1 repeat homology <THR1>

F;135-191/Domain: thrombospondin type 1 repeat homology <THR2>

F;192-255/Domain: thrombospondin type 1 repeat homology <THR3>

F;256-313/Domain: thrombospondin type 1 repeat homology <THR4>

F;314-377/Domain: thrombospondin type 1 repeat homology <THR5>

F;378-440/Domain: thrombospondin type 1 repeat homology <THR6>
F;83,86,139,142,145,196,199,260,263,321,324,382,385,388/Modified site: 2'-
mannosyl-tryptophan (Trp) #status experimental
F;428/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.1%; Score 243; DB 1; Length 469;
Best Local Similarity 39.5%; Pred. No. 1.1e-09;
Matches 45; Conservative 14; Mismatches 43; Indels 12; Gaps 4;

```
Qy      243 GGWSTWTEWSVCSASCGRGWQKRSRSCNPAPLNGGAFCEGQNVQKTACAT--LCPVDGS 300
          |||| | | || :| :| :| | :|| || | || : || | :|| | :
Db      137 GGWSGWGPWEPCSVTCSKGTTRRRACNHPAPKCGG-HCPGQAQSEACDTQQVCPTHGA 195

Qy      301 WSPWSKWSACGLDC-----THWSRECSDPAP--RNGGEECQGTDLDTNRCT 345
          |: | |: | | | ||:| | | : |: | | : | ||
Db      196 WATWGPWTPCSASCHGGPHEPKETRSRKCSAPEPSQKPPGKPCPGLAYEQRRCT 249
```

RESULT 15

S05478

properdin - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004

C;Accession: S05478

R;Goundis, D.; Reid, K.B.M.

Nature 335, 82-85, 1988

A;Title: Properdin, the terminal complement components, thrombospondin and the circumsporozoite protein of malaria parasites contain similar sequence motifs.

A;Reference number: S05478; MUID:88318954; PMID:3045564

A;Accession: S05478

A;Molecule type: mRNA

A;Residues: 1-437 <GOU>

A;Cross-references: UNIPROT:P11680; EMBL:X12905; NID:g53786; PIDN:CAA31389.1;
PID:g53787

C;Complex: a mixture of homodimers, homotrimers and homotetramers

C;Function:

A;Description: protects C3 convertase (C3bBb) from rapid inactivation

A;Pathway: complement alternate pathway

C;Superfamily: human properdin precursor; thrombospondin type 1 repeat homology

C;Keywords: complement alternate pathway; glycoprotein; homodimer; homotetramer;
homotrimer; plasma

F;45-97/Domain: thrombospondin type 1 repeat homology <THR1>

F;104-160/Domain: thrombospondin type 1 repeat homology <THR2>

F;161-224/Domain: thrombospondin type 1 repeat homology <THR3>

F;225-282/Domain: thrombospondin type 1 repeat homology <THR4>

F;283-345/Domain: thrombospondin type 1 repeat homology <THR5>

F;346-408/Domain: thrombospondin type 1 repeat homology <THR6>

F;52,55,108,111,114,165,168,229,232,290,293,350,353,356/Modified site: 2'-
mannosyl-tryptophan (Trp) #status predicted

F;366,396/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.8%; Score 229; DB 2; Length 437;
Best Local Similarity 40.4%; Pred. No. 9.9e-09;
Matches 46; Conservative 10; Mismatches 46; Indels 12; Gaps 4;

```
Qy      243 GGWSTWTEWSVCSASCGRGWQKRSRSCNPAPLNGGAFCEGQNVQKTACAT--LCPVDGS 300
          |||| | | || :| :| :| | |||| || | |: | || | || | :
```

Db 106 GGWSEWGPWGPCSVTCSKGTQIRQRVCDNPAPKCGG-HCPGEAQQSQACDTQKTCPTHGA 164

Qy 301 WSPWSKWSACGLDC-----THWRSRECS DPAPRN--GGEECQGTDL DTRNCT 345
|: | || | ||| || ||| : |: | | : : |:

Db 165 WASWGPWSPRSGSCLGGAQEPKETRSRSCSAPAPSHQPPGKPCSGPAYEHKACS 218

Search completed: March 1, 2005, 09:07:16
Job time : 38.0558 secs

OM protein - protein search, using sw model

Run on: March 1, 2005, 09:06:01 ; Search time 172.833 Seconds
(without alignments)
1704.439 Million cell updates/sec

Title: US-10-624-932-2
Perfect score: 4791
Sequence: 1 MAVRPGLWPALLGIVLAWL.....AVAGLGQPDAGLFTVSEAE 898

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1385339 seqs, 328044528 residues

Total number of hits satisfying chosen parameters: 1385339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4791	100.0	898	10	US-09-918-779-2	Sequence 2, Appli
2	4791	100.0	898	15	US-10-624-932-2	Sequence 2, Appli
3	4698.5	98.1	899	10	US-09-970-944-2	Sequence 2, Appli
4	4638	96.8	898	10	US-09-933-261-5	Sequence 5, Appli
5	4638	96.8	898	10	US-09-970-944-13	Sequence 13, Appl
6	4638	96.8	898	14	US-10-256-702-5	Sequence 5, Appli
7	4638	96.8	898	14	US-10-240-154-16	Sequence 16, Appl
8	4413	92.1	842	15	US-10-311-623-1	Sequence 1, Appli
9	2845	59.4	544	10	US-09-970-944-14	Sequence 14, Appl
10	2815.5	58.8	557	10	US-09-933-261-6	Sequence 6, Appli
11	2815.5	58.8	557	14	US-10-256-702-6	Sequence 6, Appli
12	2787	58.2	931	10	US-09-970-944-15	Sequence 15, Appl
13	2787	58.2	931	11	US-09-972-211-121	Sequence 121, App
14	2787	58.2	931	15	US-10-087-684-35	Sequence 35, Appl
15	2787	58.2	931	15	US-10-037-417-117	Sequence 117, App
16	2787	58.2	931	15	US-10-096-625-121	Sequence 121, App
17	2787	58.2	1010	15	US-10-218-779-35	Sequence 35, Appl
18	2762	57.6	931	10	US-09-970-944-16	Sequence 16, Appl
19	2762	57.6	931	11	US-09-972-211-125	Sequence 125, App
20	2762	57.6	931	15	US-10-096-625-125	Sequence 125, App
21	2759	57.6	931	15	US-10-037-417-120	Sequence 120, App
22	2755	57.5	931	10	US-09-970-944-17	Sequence 17, Appl
23	2755	57.5	931	11	US-09-972-211-122	Sequence 122, App
24	2755	57.5	931	15	US-10-087-684-36	Sequence 36, Appl
25	2755	57.5	931	15	US-10-218-779-36	Sequence 36, Appl
26	2755	57.5	931	15	US-10-037-417-118	Sequence 118, App
27	2755	57.5	931	15	US-10-037-417-119	Sequence 119, App
28	2755	57.5	931	15	US-10-096-625-122	Sequence 122, App
29	2578.5	53.8	945	11	US-09-972-211-124	Sequence 124, App
30	2578.5	53.8	945	15	US-10-087-684-34	Sequence 34, Appl
31	2578.5	53.8	945	15	US-10-218-779-34	Sequence 34, Appl
32	2578.5	53.8	945	15	US-10-037-417-121	Sequence 121, App
33	2578.5	53.8	945	15	US-10-096-625-124	Sequence 124, App
34	2572.5	53.7	945	11	US-09-972-211-123	Sequence 123, App
35	2572.5	53.7	945	15	US-10-087-684-33	Sequence 33, Appl
36	2572.5	53.7	945	15	US-10-218-779-33	Sequence 33, Appl
37	2572.5	53.7	945	15	US-10-096-625-123	Sequence 123, App
38	2571.5	53.7	943	10	US-09-933-261-7	Sequence 7, Appli
39	2571.5	53.7	943	14	US-10-256-702-7	Sequence 7, Appli
40	2563.5	53.5	933	15	US-10-087-684-2	Sequence 2, Appli
41	2563.5	53.5	933	15	US-10-087-684-4	Sequence 4, Appli
42	2563.5	53.5	933	15	US-10-218-779-2	Sequence 2, Appli
43	2563.5	53.5	933	15	US-10-218-779-4	Sequence 4, Appli
44	2558.5	53.4	945	14	US-10-028-072-146	Sequence 146, App
45	2558.5	53.4	945	14	US-10-140-808-146	Sequence 146, App

ALIGNMENTS

RESULT 1

US-09-918-779-2

; Sequence 2, Application US/09918779

; Publication No. US20030064369A1

```

; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Spytek, Kimberly
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Gusev, Vladimir
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook, John
; APPLICANT: Lepley, Denise
; APPLICANT: Burgess, Catherine
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Smithson, Glennnda
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-074 US
; CURRENT APPLICATION NUMBER: US/09/918,779
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/221,409
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/222,840
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,752
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,762
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,770
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,769
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/225,146
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,392
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,470
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,697
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/263,662
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/281,645
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-779-2

```

Query Match

100.0%; Score 4791; DB 10; Length 898;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 898; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
        |||
Db      1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGL 120
        |||
Db     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGL 120

Qy    121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE 180
        |||
Db    121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE 180

Qy    181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY 240
        |||
Db    181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY 240

Qy    241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVQKTACATLCPVDGS 300
        |||
Db    241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVQKTACATLCPVDGS 300

Qy    301 WSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVA 360
        |||
Db    301 WSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVA 360

Qy    361 LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHELL 420
        |||
Db    361 LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHELL 420

Qy    421 TIQPDLSSTTTTYYQGSCLCPQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFEVS 480
        |||
Db    421 TIQPDLSSTTTTYYQGSCLCPQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFEVS 480

Qy    481 RLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYEIYLTLHK 540
        |||
Db    481 RLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYEIYLTLHK 540

Qy    541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCEGSPDSWSLRLKKQSCGSW 600
        |||
Db    541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCEGSPDSWSLRLKKQSCGSW 600

Qy    601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLLFAPVACT 660
        |||
Db    601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLLFAPVACT 660

Qy    661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLW 720
        |||
Db    661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLW 720

Qy    721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWWVQVEGDGQSFSINF 780
        |||
Db    721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWWVQVEGDGQSFSINF 780

Qy    781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPCCRRGADWRTLAQKL 840
        |||
```

Db 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL 840

Qy 841 HLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898
 |||

Db 841 HLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898

RESULT 2

US-10-624-932-2

; Sequence 2, Application US/10624932
 ; Publication No. US20040096877A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Taupier, Raymond
 ; APPLICANT: Padigar, Muralidhara
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Spaderna, Steven
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Zerhusen, Bryan
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Shenoy, Suresh
 ; APPLICANT: Li, Li
 ; APPLICANT: Gusev, Vladimir
 ; APPLICANT: Grosse, William
 ; APPLICANT: Alsobrook, John
 ; APPLICANT: Lepley, Denise
 ; APPLICANT: Burgess, Catherine
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: MacDougall, John
 ; APPLICANT: Stone, David
 ; APPLICANT: Smithson, Glennnda
 ; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-074 US
 ; CURRENT APPLICATION NUMBER: US/10/624,932
 ; CURRENT FILING DATE: 2003-07-21
 ; PRIOR APPLICATION NUMBER: 09/918,779
 ; PRIOR FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: 60/221,409
 ; PRIOR FILING DATE: 2000-07-28
 ; PRIOR APPLICATION NUMBER: 60/222,840
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: 60/223,752
 ; PRIOR FILING DATE: 2000-08-08
 ; PRIOR APPLICATION NUMBER: 60/223,762
 ; PRIOR FILING DATE: 2000-08-08
 ; PRIOR APPLICATION NUMBER: 60/223,770
 ; PRIOR FILING DATE: 2000-08-08
 ; PRIOR APPLICATION NUMBER: 60/223,769
 ; PRIOR FILING DATE: 2000-08-08
 ; PRIOR APPLICATION NUMBER: 60/225,146
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/225,392
 ; PRIOR FILING DATE: 2000-08-15
 ; PRIOR APPLICATION NUMBER: 60/225,470
 ; PRIOR FILING DATE: 2000-08-15
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-624-932-2

Query Match 100.0%; Score 4791; DB 15; Length 898;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 898; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFIVEPEDVYIVKNKP	60
Db	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFIVEPEDVYIVKNKP	60
Qy	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL	120
Db	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL	120
Qy	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE	180
Db	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE	180
Qy	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY	240
Db	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Qy	301	WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVA	360
Db	301	WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVA	360
Qy	361	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSGFQPVSIKPSKADNPHLL	420
Db	361	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSGFQPVSIKPSKADNPHLL	420
Qy	421	TIQPDLS TTTT TYQGS LCP RQD G P S P K F Q L T N G H L L S P L G G R H T L H H S S P T S E A E E F V S	480
Db	421	TIQPDLS TTTT TYQGS LCP RQD G P S P K F Q L T N G H L L S P L G G R H T L H H S S P T S E A E E F V S	480
Qy	481	RLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLHK	540
Db	481	RLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLHK	540
Qy	541	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSW	600
Db	541	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSW	600
Qy	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACT	660
Db	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACT	660
Qy	661	SLEYNIRVYCLHDT HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW	720

[illegible]

Db	181	VEWLRNEDLVDP SLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNV-QKTACATLCPVDG	299
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVHDRTVSSLLVSVDG	300
Qy	300	SWSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDV	359
Db	301	SWSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDV	360
Qy	360	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTS GFQPVSIKPSKADNPHL	419
Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTS GFQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLS TTTTYYQGS LCPQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Db	421	LTIQPDLS -TTTYYQGS LCPQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFV	479
Qy	480	SRLSTQNYFRSLPRGTSNM TYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LH	539
Db	480	SRLSTQNYFRSLPRGTSNM TYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPS PDSWSLRLKKQSCGS	599
Db	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPS PDSWSLRLKKQSCGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVA	659
Qy	659	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS	718
Db	660	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS	719
Qy	719	LWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSI	778
Db	720	LWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSI	779
Qy	779	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	838
Db	780	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	839
Qy	839	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	898
Db	840	KLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	899

RESULT 4

US-09-933-261-5

; Sequence 5, Application US/09933261

; Publication No. US20030040046A1

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

```

; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030040046A1 Relevant
; TOPOLOGY: No. US20030040046A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-933-261-5

```

```

Query Match          96.8%; Score 4638; DB 10; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

```

```

Qy      1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL 120
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRRQVEKVFGL 120

Qy    121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPP 180
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db    121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPP 180

Qy    181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db    181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAVIVY 240

```

Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCSTNPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCSTNPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Qy	301	WSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVA	360
Db	301	WSSWSKWSACGLDCTHWSRECSDPAPRNGGEECRGADLDTNRCTSDLCVHSASGPEDVA	360
Qy	361	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSQFQPVSIKPSKADNPHLL	420
Db	361	LYIGLVAVAVCLFLLLLLALGLIYCRKKEGLDSDVADSSILTSQFQPVSIKPSKADNPHLL	420
Qy	421	TIQPDLSSTTTTQYQSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSEAEFVS	480
Db	421	TIQPDLSSTTTTQYQSLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHSSPTSEAEFVS	480
Qy	481	RLSTQNYFRSLPRGTSNMITYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHK	540
Db	481	RLSTQNYFRSLPRGTSNMAYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHK	540
Qy	541	PEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPSPDSSWSLRLKKQSCEGSW	600
Db	541	PEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPSPDSSWSLRLKKQSCEGSW	600
Qy	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAALKRLKLLLFAPVACT	660
Db	601	EDVLHLGEEAPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACT	660
Qy	661	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW	720
Db	661	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW	720
Qy	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINF	780
Db	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERINASTSDLACKLWVWQVEGDGQSFNINF	780
Qy	781	NITKDTRFAELLALAESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL	840
Db	781	NITKDTRFAELLALAESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCSRAGADWRTLAQKL	840
Qy	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	898
Db	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGOLAAAVAGLGQPDAGLFTVSEAE	898

```

; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-970-944-13

```

```

Query Match          96.8%; Score 4638; DB 10; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

```

```

Qy      1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGL 120
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSQQVEKVFGL 120

Qy    121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAE 180
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db    121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAE 180

Qy    181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY 240
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db    181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAIVIVY 240

Qy    241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVQKTACATLCPVDGS 300
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db    241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVQKTACATLCPVDGS 300

Qy    301 WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVA 360
        || ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db    301 WSSWSKWSACGLDCTHWSRECS DPAPRNGGEECRGADLDRNCTSDLCVHSASGPEDVA 360

Qy    361 LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHLL 420
        ||:|:|:| |||| | | :| ||||||| ||||||| ||||||| |||||||
Db    361 LYIGLVAVAVCLFLLLLALGLIYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHLL 420

Qy    421 TIQPDLS TTTT TYQGS LCP RQD GPSPKFQ LTN GHLLS PLGGGRHTLHHSSPTSEAEFVS 480
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db    421 TIQPDLS TTTT TYQGS LCSRQD GPSPKFQ LSN GHLLS PLGSGRHTLHHSSPTSEAEFVS 480

Qy    481 RLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLHK 540
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db    481 RLSTQNYFRSLPRGTSNMAYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYTLHK 540

Qy    541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSSWLRLLKKQSCGGSW 600
        ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db    541 PEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDSSWLRLLKKQSCGGSW 600

```

Qy	601	EDVLHLEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA	AAKRLKLLLFAPVACT	660
Db	601	EDVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALS	VAAATKRLRLLLFAPVACT	660
Qy	661	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIH	DVPSSLW	720
Db	661	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIH	DVPSSLW	720
Qy	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLVVWQVEGDG	QSFSINF	780
Db	721	KSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDG	QSFNINF	780
Qy	781	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRT	LAQKL	840
Db	781	NITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSR	GADWRTLAQKL	840
Qy	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTV	SEAC	898
Db	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGOLAAAVAGLGOPDAGLFTV	SEAC	898

```

;          REFERENCE/DOCKET NUMBER: UC96-217
;          TELECOMMUNICATION INFORMATION:
;          TELEPHONE: (415) 343-4341
;          TELEFAX: (415) 343-4342
;          INFORMATION FOR SEQ ID NO: 5:
;          SEQUENCE CHARACTERISTICS:
;          LENGTH: 898 amino acids
;          TYPE: amino acid
;          STRANDEDNESS: No. US20030059859A1 Relevant
;          TOPOLOGY: No. US20030059859A1 Relevant
;          MOLECULE TYPE: peptide
;          SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-256-702-5

```

```

Query Match          96.8%; Score 4638; DB 14; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

```

```

Qy      1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
        |||
Db      1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL 120
        |||
Db     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSRRQVEKVFGL 120

Qy    121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAAE 180
        |||
Db    121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAAE 180

Qy    181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
        |||
Db    181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAAVIVY 240

Qy    241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVQKTACATLCPVDGS 300
        |||
Db    241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVQKTACATLCPVDGS 300

Qy    301 WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVA 360
        || |||
Db    301 WSSWSKWSACGLDCTHWSRECS DPAPRNGGEECRGADLDRNCTSDLCVHSASGPEDVA 360

Qy    361 LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILTSGFQPVSIKPSKADNPHLL 420
        |||
Db    361 LYIGLVAVAVCLFLLLLLALGLIYCRKKEGLSDVDADSSILTSGFQPVSIKPSKADNPHLL 420

Qy    421 TIQPDLS TTTT TYQGS LCP RQD GSP K F Q L T N G H L L S P L G G R H T L H H S S P T S E A E E F V S 480
        |||
Db    421 TIQPDLS TTTT TYQGS LCP RQD GSP K F Q L S N G H L L S P L G S G R H T L H H S S P T S E A E D F V S 480

Qy    481 RLSTQNYFRSLPRGTSNM TYGT FN FL G G R L M I P N T G I S L L I P P D A I P R G K I Y E I Y L T L H K 540
        |||
Db    481 RLSTQNYFRSLPRGTSNM MAYGT FN FL G G R L M I P N T G I S L L I P P D A I P R G K I Y E I Y L T L H K 540

Qy    541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCGSGW 600
        |||
Db    541 PEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCGSGW 600

```


Qy	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACT	660
Db	601	EDVLHLGEEAPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACT	660
Qy	661	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW	720
Db	661	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW	720
Qy	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLVVWQVEGDGQSFSINF	780
Db	721	KSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNINF	780
Qy	781	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLQKL	840
Db	781	NITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSR GADWRTLQKL	840
Qy	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	898
Db	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGOLAAAVAGLGOPDAGLFTVSEAE	898

US-10-240-154-16

Query Match 96.8%; Score 4638; DB 14; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

Qy	181	VEWLRNEDLVDP	SLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASA	AIVY	240
Db	181	VEWLRNEDLVDP	SLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSA	AIVY	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSR	SCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS		300
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSR	SCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS		300
Qy	301	WSPWSKWSACGLDCTHWSRECS	DPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVA		360
Db	301	WSSWSKWSACGLDCTHWSRECS	DPAPRNGGEECRGADLDTRNCTSDLCVHSASGPEDVA		360
Qy	361	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLDS	SDVADSSILTSGFQPVSIKPSKADNP	PHLL	420
		: :			
Db	361	LYIGLVAVAVCLFLLLLALGLIYCRKKEGLDS	SDVADSSILTSGFQPVSIKPSKADNP	PHLL	420
Qy	421	TIQPDLS	TTTTTYQGS	SLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE	FVS 480
Db	421	TIQPDLS	TTTTTYQGS	SLCSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAE	FVS 480
Qy	481	RLSTQNYFRSLPRGTSNMTYGT	FNFLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT	HLK	540
Db	481	RLSTQNYFRSLPRGTSNMTYGT	FNFLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT	HLK	540
Qy	541	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAM	DHCGEPPDSWSLRLKKQSC	EGSW	600
Db	541	PEDVRLPLAGCQTLLSPVWSCGPPGVLLTRPVILAM	DHCGEPPDSWSLRLKKQSC	EGSW	600
Qy	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	VAAAKRLKLLLFAPVACT		660
Db	601	EDVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALS	VAAAKRLKLLLFAPVACT		660
Qy	661	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHN	LRLSIHDPSSLW		720
Db	661	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHN	LRLSIHDPSSLW		720
Qy	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSF	SINF		780
Db	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERINASTDLACKLWVWQVEGDGQSF	NINF		780
Qy	781	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKI	ISSLDPPCRRGADWRTLAQKL		840
Db	781	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKI	ISSLDPPCRRGADWRTLAQKL		840
Qy	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	C		898
Db	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	C		898

RESULT 8

US-10-311-623-1

; Sequence 1, Application US/10311623

; Publication No. US20040023244A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.

; APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.

```

; APPLICANT: YUE, Henry; NGUYEN, Danniel B.
; APPLICANT: TANG, Y. Tom; LAL, Preeti G.
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda
; APPLICANT: LU, Dyung Aina M.; GRAUL, Richard C.
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah R.
; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.
; APPLICANT: YANG, Junming; XU, Yuming
; APPLICANT: GANDHI, Ameena R.; WARREN, Bridget A.
; APPLICANT: DING, Li; SANJANWALA, Madhusudan M.
; APPLICANT: DUGGAN, Brendan M.; LU, Yan
; TITLE OF INVENTION: RECEPTORS
; FILE REFERENCE: PF-0793 USN
; CURRENT APPLICATION NUMBER: US/10/311,623
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 01/19942
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/214,027
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/228,045
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/255,104
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CD1
US-10-311-623-1

```

```

Query Match          92.1%; Score 4413; DB 15; Length 842;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 840; Conservative 1; Mismatches 1; Indels 56; Gaps 1;

```

```

Qy      1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
        |||
Db      1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGL 120
        |||
Db     61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGL 120

Qy    121 EYWCQCVAWSSSGTTSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPP 180
        |||
Db    121 EYWCQCVAWSSSGTTSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPP 180

Qy    181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASA 240
        |||
Db    181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASA 240

Qy    241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVQKTACATLCPVDGS 300
        |||
Db    241 -----VDGS 244

```

Qy	301	WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVA	360
Db	245	WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDL DTRNCTSDLCVHTASGPEDVA	304
Qy	361	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDASSILTSGFQPVSIKPSKADNPHELL	420
Db	305	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDASSILTSGFQPVSIKPSKADNPHELL	364
Qy	421	TIQPD LSTTTTTTYQGS LCP RQDGSPK FQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVS	480
Db	365	TIQPD LSTTTTTTYQGS LCP RQDGSPK FQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVS	424
Qy	481	RLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHK	540
Db	425	RLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHK	484
Qy	541	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD SWSLRLKKQSCEGSW	600
Db	485	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD SWSLRLKKQSCEGSW	544
Qy	601	EDVLHLGEEAPSHLYYCQEASAC YVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACT	660
Db	545	EDVLHLGEEAPSHLYYCQEASAC YVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACT	604
Qy	661	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEP RVLHFKDSYHNLRLSIHDVPSSLW	720
Db	605	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEP RVLHFKDSYHNLRLSIHDVPSSLW	664
Qy	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINF	780
Db	665	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINF	724
Qy	781	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKI ISSLDPPCRRGADWRTLAQKL	840
Db	725	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKI ISSLDPPCRRGADWRTLAQKL	784
Qy	841	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAA AVAGLGQPDAGLFTVSEAE C	898
Db	785	HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAA AVAGLGQPDAGLFTVSEAE C	842

RESULT 9

US-09-970-944-14

; Sequence 14, Application US/09970944

; Publication No. US20030204052A1

; GENERAL INFORMATION:

; APPLICANT: Herrman, John L

; APPLICANT: Rastelli, Luca

; APPLICANT: Shimkets, Richard A

; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same and

; TITLE OF INVENTION: Antibodies Directed Against these Proteins

; FILE REFERENCE: 21402-138

; CURRENT APPLICATION NUMBER: US/09/970,944

; CURRENT FILING DATE: 2002-05-02

; PRIOR APPLICATION NUMBER: 60/237,862

; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-14

Query Match 59.4%; Score 2845; DB 10; Length 544;
Best Local Similarity 100.0%; Pred. No. 2.7e-225;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	358	DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSQFQPVSIKPSKADNP	417
Db	4	DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSQFQPVSIKPSKADNP	63
Qy	418	HLLTIQPDLSSTTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE	477
Db	64	HLLTIQPDLSSTTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE	123
Qy	478	FVSRSLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT	537
Db	124	FVSRSLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT	183
Qy	538	LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSRLKKQSCE	597
Db	184	LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSRLKKQSCE	243
Qy	598	GSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLFAPV	657
Db	244	GSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLFAPV	303
Qy	658	ACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPS	717
Db	304	ACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPS	363
Qy	718	SLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFS	777
Db	364	SLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFS	423
Qy	778	INFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLA	837
Db	424	INFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLA	483
Qy	838	QKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	897
Db	484	QKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	543
Qy	898	C 898	
Db	544	C 544	

RESULT 10
US-09-933-261-6
; Sequence 6, Application US/09933261

; Publication No. US20030040046A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030040046A1 Relevant
; TOPOLOGY: No. US20030040046A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-933-261-6

Query Match 58.8%; Score 2815.5; DB 10; Length 557;
Best Local Similarity 96.8%; Pred. No. 7.4e-223;
Matches 539; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

```
Qy      343 NCTSDLCVHSASGPEDVALYVGLIHAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 402
          |||:|||||
Db      1 NCTSDLXVHTASGPEDVALYVGLIHAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 60

Qy      403 GFQPVSIKPSKADNPHELLTIQPDLSITTTTYYQGSILCPRQDGPSPKFQLTNGHLLSPLGGG 462
          |||:|||||
Db      61 GFQPVSIKPSKADNPHELLTIQPDLSITTTTYYQGSILCPRQDGPSPKFQLTNGHLLSPLGGG 120
```

Qy	463	RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIP	522
Db	121	RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIP	180
Qy	523	PDAIPRGKIYEIYLTLLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE	582
Db	181	PDAIPRGKIYEIYLTLLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE	240
Qy	583	SPDSWSLRLKKQSCGWSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	642
Db	241	SPDSWSLALKKQSCGWSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	300
Qy	643	VAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFK	702
Db	301	VAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHLX	360
Qy	703	DSYHNLRSLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLAC	762
Db	361	DSYHNLRSLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLAC	420
Qy	763	KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS	822
Db	421	KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS	480
Qy	823	LDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG	882
Db	481	LDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG	540
Qy	883	LGQPDAGLFT-VSEAEC	898
		:	
Db	541	TXPAGRWLLSQCEAEC	557

RESULT 11

US-10-256-702-6

; Sequence 6, Application US/10256702

; Publication No. US20030059859A1

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc
 ; Leonardo, E. David
 ; Hink, Lindsay
 ; Masu, Masayuki
 ; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 268 BUSH STREET, SUITE 3200
 ; CITY: SAN FRANCISCO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

Qy 763 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 822
 |||
 Db 421 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 480
 Qy 823 LDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG 882
 |||
 Db 481 LDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG 540
 Qy 883 LGQPDAGLFT-VSEAE 898
 | : ||||
 Db 541 TXPAGRWLLSQCSAE 557

RESULT 12

US-09-970-944-15

; Sequence 15, Application US/09970944
 ; Publication No. US20030204052A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Herrman, John L
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Shimkets, Richard A
 ; TITLE OF INVENTION: No. US20030204052A1e1 Proteins and Nucleic Acids Encoding Same and
 ; TITLE OF INVENTION: Antibodies Directed Against these Proteins
 ; FILE REFERENCE: 21402-138
 ; CURRENT APPLICATION NUMBER: US/09/970,944
 ; CURRENT FILING DATE: 2002-05-02
 ; PRIOR APPLICATION NUMBER: 60/237,862
 ; PRIOR FILING DATE: 2000-10-04
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 15
 ; LENGTH: 931
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-09-970-944-15

Query Match 58.2%; Score 2787; DB 10; Length 931;
 Best Local Similarity 57.3%; Pred. No. 3.5e-220;
 Matches 522; Conservative 153; Mismatches 208; Indels 28; Gaps 9;

Qy 9 PALLGIVLAAWLRGSGAQQA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
 ||| :|| ||| : | : ||||:|:| : ||||| |
 Db 26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSPPEPLPHFLIEPEEAYIVKNKPVNLYC 83
 Qy 66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGLLEEYWCQ 125
 || ||||:|||| ||| | |||: | :|| | | :|||:|:| | :|||
 Db 84 KASPATQIYFKCNSEWVHQDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143
 Qy 126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLR 185
 ||||:||||:|||| ||| ||||| ||||| :| ||||| ||||| :
 Db 144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203
 Qy 186 NEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
 |||:|:| | | ||| :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
 Db 204 NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy 246 STWTEWSVCSASCGRGWQKRSRSCCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWS 305
 |||||:: ||||:||||:|||||||:|||| || ||||| |: ||
 Db 264 STWTEWSVCNSRCGRGYQKRTRCTNPAPLNGGAFCEGQSVQKIACCTTLCPVDGRWTSWS 323
 Qy 306 KWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGL 365
 ||| || :||| |||: |||:||||:| | | :||| ||: |: | |||||:
 Db 324 KWSTCGTECTHWRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPSDDDVALYVGI 383
 Qy 366 -IAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHELLTIQP 424
 ||| ||| : ::| : || : ||: ||| | ||||:| | :| || : |
 Db 384 VIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP 440
 Qy 425 DLSTTTTTTYQGSCLCPQDGPSPKFQLTNGHLLSPLGGGRHTLHHSS----PTSEAEFEVS 480
 ||:: | |: | | :| :| | : ::|| | : || |
 Db 441 DLTSAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSS 499
 Qy 481 RLS---TQNYF-----RSLPRGT--SNMTYGTENFLGGRLMIPNTGISLLIPPDAI 526
 :|| ||: || | | :||| || |:| |:| || ||
 Db 500 KLSPQMTQSLENEALNLKNQSLARQTDPSCTAFGTENSLGGHLIIPNSGVSLIPAGAI 559
 Qy 527 PRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDDS 586
 |:|::| |:| |:| |:| | | |||:|:| ||||| | || :| :
 Db 560 PQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVS CGPPGALLTRPVILT LHHCADPSTED 619
 Qy 587 WSLRLKKQSCGWSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 646
 | :|| |: |:| |||: ||| : | ||:| ||:: || | :|||:: : |||
 Db 620 WKIQLKNQAVQGWEDVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAA 679
 Qy 647 KRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH 706
 |||| |: |: |:| |||:| |||| || |||||:| |:| |:| |:| | |
 Db 680 KRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEPKALHFKGSIH 739
 Qy 707 NLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWV 766
 |||||: ||||| |||||:|:| |||||:| |:| |:| ||| |
 Db 740 NLRLSIHDIAHSLWKS KLLAKYQEIPFYHIWSGSQRNHCTFTLERLSLNTVELVCKLCV 799
 Qy 767 WQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKISSLDP 826
 ||||:| | |:| :::: |:| |:| : : |||| || |||: |||| |
 Db 800 RQVEGEGQIFQLNCTVSEEP TGIDPLLDPASTITVTGPSAFSIPPIRQKLCSSLDAP 859
 Qy 827 CRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQP 886
 || ||| || ||:| |:|:| |:| || |:|:| |||:| |:| ||| ||| : |:|
 Db 860 QTRGHDWRMLAHKLNLD RYLN YFATKSSPTGVILDLWEAQNFDPGNLSMLAAVLEEMGRH 919
 Qy 887 DAGLFTVSEAE 897
 : : :| :
 Db 920 ETVVSLAAEGQ 930

RESULT 13

US-09-972-211-121

; Sequence 121, Application US/09972211

; Publication No. US20040048245A1

; GENERAL INFORMATION:

; APPLICANT: Shimkets, Richard A

; APPLICANT: Taupier Jr, Raymond J

; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennda
; APPLICANT: Szekeres Jr, Edward S
; TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides
Encoding Them And
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141
; CURRENT APPLICATION NUMBER: US/09/972,211
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 931

; TYPE: PRT
; ORGANISM: Mus musculus
US-09-972-211-121

Query Match 58.2%; Score 2787; DB 11; Length 931;
Best Local Similarity 57.3%; Pred. No. 3.5e-220;
Matches 522; Conservative 153; Mismatches 208; Indels 28; Gaps 9;

Qy	9	PALLGIVLAAWLRGSGAQQSA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC	65
		: : :	
Db	26	PAL--ALLSASGTGSAAQDDEFFHELPETFPSPPEPLPHFLIEPEEAYIVKNKPVNLYC	83
Qy	66	KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQ	125
		: :	
Db	84	KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLIVREVSIERSRQVEELFGPEDYWCQ	143
Qy	126	CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLR	185
		: :	
Db	144	CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPGEGIPVAEVEWLK	203
Qy	186	NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW	245
		: : : : : :	
Db	204	NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW	263
Qy	246	STWTEWSVCSASCGRGWQKRSRCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWS	305
		: : :	
Db	264	STWTEWSVCNSRCGRGYQKRTRCTNPAPLNGGAFCEGQSVQKIACCTTLCPCVDGRWTSWS	323
Qy	306	KWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGL	365
		: : : : : :	
Db	324	KWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPSDDDVALYVGI	383
Qy	366	-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHELLTIQP	424
		: : : :	
Db	384	VIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP	440
Qy	425	DLSTTTTTTYQGSCLPRQDGSPKPFQLTNGHLLSPLGGGRHTLHHSS-----PTSEAEFVS	480
		: : : : : : : : :	
Db	441	DLTSAAAMYRGVPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSS	499
Qy	481	RLS---TQNYF-----RSLPRGT--SNMTYGTFFNLGGRLMIPNTGISLLIPDAI	526
		: : : :	
Db	500	KLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFFNSLGGHLIIPNSGVSLIPAGAI	559
Qy	527	PRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPS	586
		: : : : : : :	
Db	560	PQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVWSCGPPGALLTRPVILT LHHCADPSTED	619
Qy	587	WSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA	646
		: : : : : :	
Db	620	WKIQLKNQAVQGQWEDVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAA	679
Qy	647	KRLKLLL FAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH	706
		: :	
Db	680	KRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGSIH	739
Qy	707	NLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWV	766

Db	740	NLRLSIHDIAHSLWKSLLAKYQEIPFYHIWGSQRNLHCTFTLERLSLNTVELVCKLCV	799
Qy	767	WQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKISSLDPP	826
Db	800	RQVEGEGQIFQLNCTVSEPTGIDLPLLDPASTITVTGSPAFSIPLPIRQKLCSSLDAP	859
Qy	827	CRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQP	886
Db	860	QTRGHDWRMLAHKLNLDRLNYFATKSSPTGVILDLWEAQNFDPGNLSMLAAVLEEMGRH	919
Qy	887	DAGLFTVSEAE	897
Db	920	ETVVSLLAEGQ	930

RESULT 14

US-10-087-684-35

; Sequence 35, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Cathereine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08

```
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 35
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-684-35
```

```
Query Match          58.2%; Score 2787; DB 15; Length 931;
Best Local Similarity 57.3%; Pred. No. 3.5e-220;
Matches 522; Conservative 153; Mismatches 208; Indels 28; Gaps 9;
```

```
Qy      9 PALLGIVLAAWLRGSGAQQA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
      |||  :|:|  ||  |  :  |  :|  |||:|:|:|  |||  |||
Db     26 PAL--ALLSASGTGSAAQDDEFFHELPEFSPDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy     66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEWYCQ 125
      ||  |||:|:|  |||  |  |||:  |  :||  ||  |  :|:|:|:|:|  |  :|:|
Db     84 KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLVREVSIERSRQQVEELFGPEDYWCQ 143

Qy    126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLR 185
      |||:|:|:|:|:|:|:|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db    144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy    186 NEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVYVNGGW 245
      |||:|:|  |  |  |||  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  |  :|  |||
Db    204 NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy    246 STWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWS 305
      |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  ||  |||  |||  |  :  ||
Db    264 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACCTTLCPVDGRWTSWS 323

Qy    306 KWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVALYVGL 365
      |||  ||  :|:|  |||:  |||:|:|:|  |  |  :|:|  ||:  :|  :|:|:|:|
Db    324 KWSTCGTECTHWRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPSDDDVALYVGI 383

Qy    366 -IAVAVCLVLLLVLILVYCRKKEGLSDVADSSILTSQFQPVSIKPSKADNPHLLTIQP 424
      |||  |||  :  :|  :  ||  :  :|:  |||  |  |||:|  :|  ||  :  |
Db    384 VIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP 440

Qy    425 DLSTTTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSS----PTSEAEFEVS 480
      ||:  :  |  :  |  |  |  :|  :|  ||  :  :|:|  |  :  |||
Db    441 DLTSAAMYRGFPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSS 499

Qy    481 RLS---TQNYF-----RSLPRGT--SNMTYGTFFNLGGRLMIPNTGISLLIPDAI 526
      :||  ||:  :|  |  |  |  :|  |||  |||  |  :|:|:|:|  ||
Db    500 KLSQMTOQLLENEALNLKNQSLARQTDPSCTAFGTFNLSLGGHLIIPNSGVSLLIPAGAI 559

Qy    527 PRGKIYEIYLTLHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSPDS 586
      |:|:|:|:|:|:|:|  |:|  |  |  |||:|:|:|  |||  |||  :  ||  :|  :
Db    560 PQGRVYEMYVTVHRKENMRPPMEDSQTLTPVVSCTGPPGALLTRPVILTLLHHCADPSTED 619

Qy    587 WSLRLKKQSCGWSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 646
```

Db	620	WKIQLKNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAA	679
Qy	647	KRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH	706
Db	680	KRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGSIH	739
Qy	707	NLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWV	766
Db	740	NLRLSIHDIHSLWKSLLAKYQEIPFYHIWSGSQRNHCTFTLERLSLNTVELVCKLCV	799
Qy	767	WQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDP	826
Db	800	RQVEGEGQIFQLNCTVSEETPIDLPLLDPASTITVTGPSAFSIPPLIRQKLCSSLDAP	859
Qy	827	CRRGADWRTLAQKLHLDLHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQP	886
Db	860	QTRGHDWRMLAHKLNLDRLNYFATKSSPTGVILDLWEAQNFDPGNLSMLAAVLEEMGRH	919
Qy	887	DAGLFTVSEAE	897
Db	920	ETVVSLLAEGQ	930

RESULT 15

US-10-037-417-117

; Sequence 117, Application US/10037417

; Publication No. US20040052806A1

; GENERAL INFORMATION:

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Alsobrook II, John P

; APPLICANT: Tchernev, Velizar T

; APPLICANT: Liu, Xiaohong

; APPLICANT: Spytek, Kimberly A

; APPLICANT: Patturajan, Meera

; APPLICANT: Grosse, William M

; APPLICANT: Lepley, Denise M

; APPLICANT: Burgess, Catherine E

; APPLICANT: Vernet, Corine A.M.

; APPLICANT: Li, Li

; APPLICANT: Gorman, Linda

; APPLICANT: Edinger, Shlomit R

; APPLICANT: Sciore, Paul

; APPLICANT: Ellerman, Karen

; APPLICANT: Malyankar, Uriel M

; APPLICANT: Rothenberg, Mark

; APPLICANT: Stone, David J

; APPLICANT: Boldog, Ferenc L

; APPLICANT: Guo, Xiaojia

; APPLICANT: Shenoy, Suresh G

; APPLICANT: Anderson, David W

; APPLICANT: Padigar, Muralidhara

; APPLICANT: Taupier Jr, Raymond J

; APPLICANT: Miller, Charles E

; APPLICANT: Eisen, Andrew J

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-235

```

; CURRENT APPLICATION NUMBER: US/10/037,417
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/260,018
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,360
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/272,411
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/272,817
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/291,186
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/303,231
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/305,060
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/318,405
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/318,700
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 117
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-037-417-117

```

```

Query Match          58.2%; Score 2787; DB 15; Length 931;
Best Local Similarity 57.3%; Pred. No. 3.5e-220;
Matches 522; Conservative 153; Mismatches 208; Indels 28; Gaps 9;

```

```

Qy      9 PALLGIVLAAWLRGSGAQQSA---TVANPVPGANPDLLPHFLVEPEDVYIVKKNKPVLLVC 65
      ||| :||| ||| : | :| |||||:||||| |||
Db     26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSPDPPEPLPHFLIEPEEAYIVKKNKPVNLYC 83

Qy     66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGLLEEYWCQ 125
      || |||||:|||| ||| | |||: | :||| || | :|||||:| ||| :|||
Db     84 KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLIVREVSI EISRQQVEELFGPEDYWCQ 143

Qy    126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLR 185
      |||||:|||||:||||:||| ||| ||||| ||||| :| ||||| |||||:
Db    144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPGEGIPVAEVEWLK 203

Qy    186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASA AIVYVNGGW 245
      |||:| |: | ||| :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db    204 NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy    246 STWTEWSVCSASCGRGWQKRSRSCCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWS 305
      |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:| ||| ||| ||| :| |||
Db    264 STWTEWSVCNSRCGRGYQKRTRCTNPAPLNGGAFCEGQSVQKIACTTLCVPDGRWTSWS 323

Qy    306 KWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVALYVGL 365
      ||| || :||| |||: |||:|:|:| | | :||| ||: :| :|||:|:|:|
Db    324 KWSTCGTECTHWRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPSDDDVALYVGI 383

Qy    366 -IAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTSQFQVPSIKPSKADNPHELLTIQP 424

```


Db	384	VIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP	440
Qy	425	DLSTTTTTTYQGSCLPRQDGSPKFLQLTNGHLLSPLGGGRHTLHHSS----PTSEAEFEVS	480
Db	441	DLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSS	499
Qy	481	RLS---TQNYF-----RSLPRGT--SNMTYGTFFNLGGRLMIPNTGISLLIPDAI	526
Db	500	KLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFFNSLGGHLIIPNSGVSLIPAGAI	559
Qy	527	PRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDPS	586
Db	560	PQGRVYEMYVTVHRKENMRPPMEDSQTLTPVWSCGPPGALLTRPVILTLLHHCADPSTED	619
Qy	587	WSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA	646
Db	620	WKIQLKNQAVQGQWEDVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAA	679
Qy	647	KRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH	706
Db	680	KRLKLAIFGPLCCSSLEYSIRVYCLDDTDALKEVLQLERQMGGQLLEEPKALHFKGSIH	739
Qy	707	NLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWV	766
Db	740	NLRLSIHDIAHSLWKSLLAKYQEIPFYHIWSGSQRLHCTFTLERLSLNTVELVCKLCV	799
Qy	767	WQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPP	826
Db	800	RQVEGEGQIFQLNCTVSEEPGTIDPLPLDPASTITTVTGPSAFSIPLPIRQKLCSSLDAP	859
Qy	827	CRRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQP	886
Db	860	QTRGHDWRMLAHKLNLDRLNYFATKSSPTGVILDLEAQNFDPGNLSMLAAVLEEMGRH	919
Qy	887	DAGLFTVSEAE	897
Db	920	ETVVSLEAEGQ	930

Search completed: March 1, 2005, 09:51:25
Job time : 176.833 secs

OM protein - protein search, using sw model

Run on: March 1, 2005, 08:41:47 ; Search time 164.273 Seconds
(without alignments)
2799.282 Million cell updates/sec

Title: US-10-624-932-2
Perfect score: 4791
Sequence: 1 MAVRPGLWPALLGIVLAOWL.....AVAGLGQPDAGLFTVSEAE 898

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	4685	97.8	898	1	UN5A_MOUSE	Q8kls4 mus musculu
2	4638	96.8	898	1	UN5A_RAT	O08721 rattus norv
3	4410	92.0	842	1	UN5A_HUMAN	Q6zn44 homo sapien
4	2787	58.2	931	1	UN5C_MOUSE	O08747 mus musculu
5	2770	57.8	931	1	UN5C_RAT	Q76lx5 rattus norv
6	2761	57.6	931	1	UN5C_CHICK	Q7t2z5 gallus gall
7	2755	57.5	931	1	UN5C_HUMAN	O95185 homo sapien
8	2646.5	55.2	943	1	UN5B_XENLA	Q8jgt4 xenopus lae
9	2578.5	53.8	945	1	UN5B_MOUSE	Q8kls3 mus musculu
10	2578.5	53.8	945	1	UN5B_RAT	O08722 rattus norv
11	2558.5	53.4	945	1	UN5B_HUMAN	Q8izj1 homo sapien
12	2200	45.9	956	1	UN5D_MOUSE	Q8kls2 mus musculu
13	2199.5	45.9	953	1	UN5D_HUMAN	Q6uxz4 homo sapien
14	1050.5	21.9	876	2	Q7PW78	Q7pw78 anopheles g
15	992	20.7	1072	1	UNC5_DROME	Q95tu8 drosophila

16	977	20.4	919	1	UNC5_CAEEL	Q26261	caenorhabdi
17	891.5	18.6	759	2	Q7PW77	Q7pw77	anopheles g
18	692	14.4	199	1	UNC5_PETMA	Q9pvd5	petromyzon
19	377.5	7.9	2673	2	Q96SC3	Q96sc3	homo sapien
20	377.5	7.9	5636	2	Q96RW7	Q96rw7	homo sapien
21	370.5	7.7	1244	2	Q69YJ3	Q69yj3	homo sapien
22	333	7.0	1388	2	Q7QKD0	Q7qkd0	anopheles g
23	318	6.6	325	2	Q8I1K1	Q8ilk1	drosophila
24	300	6.3	518	2	Q8IV45	Q8iv45	homo sapien
25	298.5	6.2	1584	1	BAI1_HUMAN	O14514	homo sapien
26	296.5	6.2	1172	1	TSP2_HUMAN	P35442	homo sapien
27	293	6.1	1074	1	SM5A_HUMAN	Q13591	homo sapien
28	293	6.1	1172	1	TSP2_MOUSE	Q03350	mus musculu
29	293	6.1	1172	2	Q7TMT3	Q7tmt3	mus musculu
30	293	6.1	1172	2	Q8CG21	Q8cg21	mus musculu
31	292	6.1	1582	2	Q8CGM0	Q8cgm0	mus musculu
32	291.5	6.1	1170	1	TSP2_BOVIN	Q95116	bos taurus
33	291	6.1	1077	1	SM5A_MOUSE	Q62217	mus musculu
34	290	6.1	1173	1	TSP1_XENLA	P35448	xenopus lae
35	288.5	6.0	1088	2	Q6PCK8	Q6pck8	xenopus lae
36	286	6.0	1081	2	Q9U631	Q9u631	drosophila
37	285	5.9	1091	2	Q7YU67	Q7yu67	drosophila
38	285	5.9	1093	2	Q9VTT0	Q9vtt0	drosophila
39	284	5.9	518	2	Q6R653	Q6r653	mus musculu
40	276	5.8	1093	1	SM5B_HUMAN	Q9p283	homo sapien
41	276	5.8	1151	2	Q6DD89	Q6dd89	homo sapien
42	276	5.8	1461	2	Q8MYA8	Q8mya8	caenorhabdi
43	275.5	5.8	632	2	Q6ZPQ8	Q6zpq8	mus musculu
44	275.5	5.8	1093	1	SM5B_MOUSE	Q60519	mus musculu
45	275.5	5.8	1122	2	Q7TT33	Q7tt33	mus musculu

ALIGNMENTS

RESULT 1

UN5A_MOUSE

ID UN5A_MOUSE STANDARD; PRT; 898 AA.
AC Q8K1S4; Q6PEF7; Q80T71;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Netrin receptor UNC5A precursor (Unc-5 homolog A) (Unc-5 homolog 1).
GN Name=Unc5a; Synonyms=Kiaa1976, Unc5h1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RX MEDLINE=22239710; PubMed=12351186; DOI=10.1016/S0925-4773(02)00248-4;
RA Engelkamp D.;
RT "Cloning of three mouse unc-5 genes and their expression patterns at
RT mid-gestation.";
RL Mech. Dev. 118:191-197(2002).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 3).

RC TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-48(2003).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates
 CC axon repulsion of neuronal growth cones in the developing nervous
 CC system upon ligand binding. Axon repulsion in growth cones may be
 CC caused by its association with DCC that may trigger signaling for
 CC repulsion. It also acts as a dependence receptor required for
 CC apoptosis induction when not associated with netrin ligand (By
 CC similarity).
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC. Interacts
 CC with MAGED1. Interacts with PRKCABP, possibly mediating some
 CC interaction with PKC (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. The interaction
 CC with PRKCABP regulates its surface expression and leads to its
 CC removal from surface of neurons and growth cones (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q8K1S4-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8K1S4-2; Sequence=VSP_011697;
 CC Note=No experimental confirmation available;
 CC Name=3;
 CC IsoId=Q8K1S4-3; Sequence=VSP_011696;
 CC Note=No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Restricted to central nervous system.

CC -!- DOMAIN: The ZU5 domain mediates the interaction with MAGED1, which
 CC participates in the induction of apoptosis (By similarity).
 CC -!- PTM: Phosphorylated by PKC in vitro. Phosphorylated on cytoplasmic
 CC tyrosine residues (By similarity).
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
 CC cleavage does not take place when the receptor is associated with
 CC netrin ligand. Its cleavage by caspases is required to induce
 CC apoptosis (By similarity).
 CC -!- SIMILARITY: Belongs to the UNC-5 family.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 ZU5 domain.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AJ487852; CAD32250.1; -.
 DR EMBL; AK122575; BAC65857.1; ALT_INIT.
 DR EMBL; BC058084; AAH58084.1; -.
 DR HSSP; P07996; 1LSL.
 DR MGD; MGI:894682; Unc5a.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR011029; DEATH_like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; Death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; TSP_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Alternative splicing; Apoptosis; Developmental protein;
 KW Immunoglobulin domain; Phosphorylation; Receptor; Repeat; Signal;
 KW Transmembrane.
 FT SIGNAL 1 25 Potential.
 FT CHAIN 26 898 Netrin receptor UNC5A.
 FT DOMAIN 26 361 Extracellular (Potential).
 FT TRANSMEM 362 382 Potential.
 FT DOMAIN 383 898 Cytoplasmic (Potential).
 FT DOMAIN 44 141 Ig-like.
 FT DOMAIN 155 234 Ig-like C2-type.

FT	DOMAIN	242	296	TSP type-1 1.
FT	DOMAIN	298	350	TSP type-1 2.
FT	DOMAIN	495	598	ZU5.
FT	DOMAIN	817	897	Death.
FT	SITE	396	397	Cleavage (by caspase-3) (By similarity).
FT	SITE	661	679	Interaction with DCC (By similarity).
FT	DISULFID	65	124	By similarity.
FT	DISULFID	170	221	By similarity.
FT	CARBOHYD	107	107	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	218	218	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	343	343	N-linked (GlcNAc. . .) (Potential).
FT	VARSPLIC	1	790	Missing (in isoform 3).
FT				/FTId=VSP_011696.
FT	VARSPLIC	241	296	Missing (in isoform 2).
FT				/FTId=VSP_011697.
FT	CONFLICT	217	217	A -> P (in Ref. 3).
SQ	SEQUENCE	898 AA;	98856 MW;	59F04BA2E196C1DB CRC64;

Query Match 97.8%; Score 4685; DB 1; Length 898;
 Best Local Similarity 96.7%; Pred. No. 0;
 Matches 868; Conservative 19; Mismatches 11; Indels 0; Gaps 0;

Qy	1	MAVRPGLWPALLGIVLA	AWLRGSGAQQSATVANP	VPGANPDLLPHFLVEPE	DVYIVKNKP	60
Db	1	MAVRPGLWPALLGIVLT	AWLRGSGAQQSATVANP	VPGANPDLLPHFLVEPE	DVYIVKNKP	60
Qy	61	VLLVCKAVPATQIFFKCN	GEWVRQVDHVI	ERSTDGSSGLPTMEVR	INVS	120
Db	61	VLLVCKAVPATQIFFKCN	GEWVRQVDHVI	ERSTDGSSGLPTMEVR	INVS	120
Qy	121	EYWCQCVAWSSSGTTKS	QKAYIRIARLRKN	FEQEPLAKEVSLEQG	IVLPCR	180
Db	121	EYWCQCVAWSSSGTTKS	QKAYIRIAYLRKN	FEQEPLAKEVSLEQG	IVLPCR	180
Qy	181	VEWLRNEDLVDPSLDPN	VYITREHSLVVRQ	ARLADTANYTCVAKN	IVARRRS	240
Db	181	VEWLRNEDLVDPSLDPN	VYITREHSLVVRQ	ARLADTANYTCVAKN	IVARRRS	240
Qy	241	VNGGWSTWTEWSVCSAS	CGRGWQKRSRSC	TNPAPLNGGAFCEGQ	NVQKTAC	300
Db	241	VNGGWSTWTEWSVCSAS	CGRGWQKRSRSC	TNPAPLNGGAFCEGQ	NVQKTAC	300
Qy	301	WSPWSKWSACGLDCTH	WRSRECS	DPAPRNGGEECQGT	DLDT	360
Db	301	WSPWSKWSACGLDCTH	WRSRECS	DPAPRNGGEECRG	ADLDT	360
Qy	361	LYVGLIAVAVCLVLL	LVLLVLI	LYCRKKEGLDS	SDVADSS	420
Db	361	LYIGLVAVAVCLL	LVLLVLI	YCRKKEGLDS	SDVADSS	420
Qy	421	TIQPDLS	TTTTTYQGS	LCPRQDG	PSPKFQL	480
Db	421	TIQPDLS	TTTTTYQGS	LCPRQDG	PSPKFQL	480
Qy	481	RLSTQNYFRSLPRGT	SNMTYGT	FNFLGGRLMI	PNTGIS	540
Db	481	RLSTQNYFRSLPRGT	SNMAYGT	FNFLGGRLMI	PNTGIS	540

Qy 541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPDSWSLRLKKQSCEGSW 600
 |||
 Db 541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPDSWSLRLKKQSCEGSW 600

Qy 601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAALKRLKLLLFAPVACT 660
 |||:||||| |||:|||||
 Db 601 EDVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACT 660

Qy 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLW 720
 |||
 Db 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLW 720

Qy 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINF 780
 |||:|||||: |||: |||
 Db 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVNSTSDLACKVWVWQVEGDGQSFNINF 780

Qy 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPCCRRGADWRTLAQKL 840
 |||:|||||: |||: |||
 Db 781 NITKDTRFAEMLALESEGGVPALVGPSAFKIPFLIRQKIITSLDPPCSRGAADWRTLAQKL 840

Qy 841 HLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAE 898
 |||
 Db 841 HLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAE 898

RESULT 2

UN5A_RAT

ID UN5A_RAT STANDARD; PRT; 898 AA.
 AC 008721;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Netrin receptor UNC5A precursor (Unc-5 homolog A) (Unc-5 homolog 1).
 GN Name=Unc5a; Synonyms=Unc5hl;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
 RP SPECIFICITY.
 RC TISSUE=Ventral spinal cord;
 RX MEDLINE=97271897; PubMed=9126742;
 RA Leonardo E.D., Hinck L., Masu M., Keino-Masu K., Ackerman S.L.,
 RA Tessier-Lavigne M.;
 RT "Vertebrate homologues of C. elegans UNC-5 are candidate netrin
 RT receptors.";
 RL Nature 386:833-838(1997).
 RN [2]
 RP FUNCTION, AND INTERACTION WITH DCC.
 RX PubMed=10399920;
 RA Hong K., Hinck L., Nishiyama M., Poo M.-M., Tessier-Lavigne M.,
 RA Stein E.;
 RT "A ligand-gated association between cytoplasmic domains of UNC5 and
 RT DCC family receptors converts netrin-induced growth cone attraction to
 RT repulsion.";

RL Cell 97:927-941(1999).
 RN [3]
 RP TISSUE SPECIFICITY.
 RX PubMed=11472849;
 RA Barrett C., Guthrie S.;
 RT "Expression patterns of the netrin receptor UNC5H1 among developing
 RT motor neurons in the embryonic rat hindbrain.";
 RL Mech. Dev. 106:163-166(2001).
 RN [4]
 RP FUNCTION.
 RX PubMed=11387206; DOI=10.1093/emboj/20.11.2715;
 RA Llambi F., Causeret F., Bloch-Gallego E., Mehlen P.;
 RT "Netrin-1 acts as a survival factor via its receptors UNC5H and DCC.";
 RL EMBO J. 20:2715-2722(2001).
 RN [5]
 RP FUNCTION, SUBCELLULAR LOCATION, AND INTERACTION WITH MAGED1.
 RX PubMed=12598531; DOI=10.1074/jbc.M300415200;
 RA Williams M.E., Strickland P., Watanabe K., Hinck L.;
 RT "UNC5H1 induces apoptosis via its juxtamembrane region through an
 RT interaction with NRAGE.";
 RL J. Biol. Chem. 278:17483-17490(2003).
 RN [6]
 RP INTERACTION WITH PRKCABP, PHOSPHORYLATION, AND MUTAGENESIS OF
 RP 896-ALA--CYS-898.
 RX PubMed=14672991; DOI=23/36/11279;
 RA Williams M.E., Wu S.C.-Y., McKenna W.L., Hinck L.;
 RT "Surface expression of the netrin receptor UNC5H1 is regulated through
 RT a protein kinase C-interacting protein/protein kinase-dependent
 RT mechanism.";
 RL J. Neurosci. 23:11279-11288(2003).
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates
 CC axon repulsion of neuronal growth cones in the developing nervous
 CC system upon ligand binding. Axon repulsion in growth cones may be
 CC caused by its association with DCC that may trigger signaling for
 CC repulsion. It also acts as a dependence receptor required for
 CC apoptosis induction when not associated with netrin ligand.
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC. Interacts
 CC with MAGED1. Interacts with PRKCABP, possibly mediating some
 CC interaction with PKC.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. The interaction
 CC with PRKCABP regulates its surface expression and leads to its
 CC removal from surface of neurons and growth cones.
 CC -!- TISSUE SPECIFICITY: Mainly expressed in regions of differentiating
 CC neurons. Expressed at early stages of neural tube development in
 CC the ventral spinal cord. In developing hindbrain, it colocalizes
 CC with a number of cranial motor neuron subpopulations from
 CC embryonic E11 to E14, while DCC is expressed by motor neurons at
 CC E12. Also expressed in non-neural structures, such as the basal
 CC plane of the hindbrain and midbrain, in the developing
 CC hypothalamus, thalamus and in the pallidum.
 CC -!- DOMAIN: The ZU5 domain mediates the interaction with MAGED1, which
 CC participates in the induction of apoptosis.
 CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By
 CC similarity). Phosphorylated by PKC in vitro.
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
 CC cleavage does not take place when the receptor is associated with
 CC netrin ligand. Its cleavage by caspases is required to induce


```

CC      apoptosis.
CC      -!- SIMILARITY: Belongs to the UNC-5 family.
CC      -!- SIMILARITY: Contains 1 death domain.
CC      -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC      -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC      -!- SIMILARITY: Contains 2 TSP type-1 domains.
CC      -!- SIMILARITY: Contains 1 ZU5 domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U87305; AAB57678.1; -.
DR      HSSP; P07996; 1LSL.
DR      RGD; 621755; Unc5h1.
DR      InterPro; IPR000488; Death.
DR      InterPro; IPR011029; DEATH_like.
DR      InterPro; IPR003599; Ig.
DR      InterPro; IPR007110; Ig-like.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR008085; TSP_1.
DR      InterPro; IPR000906; ZU5.
DR      Pfam; PF00531; Death; 1.
DR      Pfam; PF00047; ig; 1.
DR      Pfam; PF00090; TSP_1; 2.
DR      Pfam; PF00791; ZU5; 1.
DR      PRINTS; PR01705; TSP1REPEAT.
DR      SMART; SM00005; DEATH; 1.
DR      SMART; SM00409; IG; 1.
DR      SMART; SM00209; TSP1; 2.
DR      SMART; SM00218; ZU5; 1.
DR      PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
DR      PROSITE; PS50835; IG_LIKE; 1.
DR      PROSITE; PS50092; TSP1; 2.
KW      Apoptosis; Developmental protein; Immunoglobulin domain;
KW      Phosphorylation; Receptor; Repeat; Signal; Transmembrane.
FT      SIGNAL          1      25      Potential.
FT      CHAIN           26     898      Netrin receptor UNC5A.
FT      DOMAIN          26     361      Extracellular (Potential).
FT      TRANSMEM        362     382      Potential.
FT      DOMAIN          383     898      Cytoplasmic (Potential).
FT      DOMAIN          44     141      Ig-like.
FT      DOMAIN          155     238      Ig-like C2-type.
FT      DOMAIN          242     296      TSP type-1 1.
FT      DOMAIN          298     350      TSP type-1 2.
FT      DOMAIN          495     598      ZU5.
FT      DOMAIN          817     897      Death.
FT      SITE            396     397      Cleavage (by caspase-3) (By similarity).
FT      SITE            661     679      Interaction with DCC (By similarity).
FT      DISULFID         65     124      By similarity.
FT      DISULFID        170     221      By similarity.
FT      CARBOHYD         107     107      N-linked (GlcNAc . . .) (Potential).
FT      CARBOHYD        218     218      N-linked (GlcNAc . . .) (Potential).

```

FT CARBOHYD 343 343 N-linked (GlcNAc. . .) (Potential).
 FT MUTAGEN 896 898 Missing: Abolishes interaction with
 FT PRKCABP.
 SQ SEQUENCE 898 AA; 98840 MW; 7A3CBCB9E7ACA135 CRC64;

Query Match 96.8%; Score 4638; DB 1; Length 898;
 Best Local Similarity 96.0%; Pred. No. 0;
 Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

Qy	1	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Db	1	MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Qy	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGL	120
Db	61	VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTMEVRINVSQQVEKVFGL	120
Qy	121	EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE	180
Db	121	EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE	180
Qy	181	VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY	240
Db	181	VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAIVIVY	240
Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Qy	301	WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVA	360
		: : : : : : : : : : : : :	
Db	301	WSSWSKWSACGLDCTHWSRECS DPAPRNGGEECRGADLDRNCTSDLCVHSASGPEDVA	360
Qy	361	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHELL	420
		: : : : : : : : : : : : :	
Db	361	LYIGLVAVAVCLFLLLLALGLIYCRKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHELL	420
Qy	421	TIQPDLSSTTTTYYQGSLSRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVS	480
		: : : : : : : : : : : :	
Db	421	TIQPDLSSTTTTYYQGSLSRQDGPSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEFVS	480
Qy	481	RLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHK	540
Db	481	RLSTQNYFRSLPRGTSNMAYGTNFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHK	540
Qy	541	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPDSWSLRLKKQSCGWS	600
		: : : : : : : : : : :	
Db	541	PEDVRLPLAGCQTLLSPVSCGPPGVLLTRPVILAMDHCGEPPDSWSLRLKKQSCGWS	600
Qy	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACT	660
		: : : : : : : : : : :	
Db	601	EDVLHLGEEAPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACT	660
Qy	661	SLEYNIRVYCLHDT HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW	720
Db	661	SLEYNIRVYCLHDT HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW	720

Qy 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINF 780
 |||:|||||:|
 Db 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERINASTDLACKVWVWQVEGDGQSFNINF 780
 |||:|||||:|
 Qy 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRGRADWRTLAQKL 840
 |||:|||||:|
 Db 781 NITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCRGRADWRTLAQKL 840
 |||:|||||:|
 Qy 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898
 |||:|||||:|
 Db 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAE 898
 |||:|||||:|

RESULT 3

UN5A_HUMAN

ID UN5A_HUMAN STANDARD; PRT; 842 AA.
 AC Q6ZN44; Q8TF26; Q96GP4;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Netrin receptor UNC5A precursor (Unc-5 homolog A) (Unc-5 homolog 1).
 GN Name=UNC5A; Synonyms=KIAA1976, UNC5H1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX PubMed=14702039; DOI=10.1038/ngl285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,

RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 302-842 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 624-728 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=21842142; PubMed=11853319;
 RA Nagase T., Kikuno R., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XXII.
 RT The complete sequences of 50 new cDNA clones which code for large
 RT proteins.";
 RL DNA Res. 8:319-327(2001).
 RN [4]
 RP INDUCTION.
 RX PubMed=12598906; DOI=10.1038/ncb943;
 RA Tanikawa C., Matsuda K., Fukuda S., Nakamura Y., Arakawa H.;
 RT "p53RDL1 regulates of p53-dependent apoptosis.";
 RL Nat. Cell Biol. 5:216-223(2003).
 RN [5]
 RP DOWN-REGULATION IN CANCER.
 RX PubMed=12655055; DOI=10.1073/pnas.0738063100;
 RA Thiebault K., Mazelin L., Pays L., Llambi F., Joly M.-O.,
 RA Scoazec J.-Y., Saurin J.-C., Romeo G., Mehlen P.;
 RT "The netrin-1 receptors UNC5H are putative tumor suppressors
 RT controlling cell death commitment.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:4173-4178(2003).
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates
 CC axon repulsion of neuronal growth cones in the developing nervous
 CC system upon ligand binding. Axon repulsion in growth cones may be
 CC caused by its association with DCC that may trigger signaling for
 CC repulsion. It also acts as a dependence receptor required for
 CC apoptosis induction when not associated with netrin ligand.

CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC. Interacts
CC with MAGED1. Interacts with PRKCABP, possibly mediating some
CC interaction with PKC (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. The interaction
CC with PRKCABP regulates its surface expression and leads to its
CC removal from surface of neurons and growth cones (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q6ZN44-1; Sequence=Displayed;
CC Note=No experimental confirmation available;
CC Name=2;
CC IsoId=Q6ZN44-2; Sequence=VSP_011694, VSP_011695;
CC Note=No experimental confirmation available;
CC Name=3;
CC IsoId=Q6ZN44-3; Sequence=VSP_011693;
CC Note=No experimental confirmation available;
CC -!- INDUCTION: By p53/TP53.
CC -!- DOMAIN: The ZU5 domain mediates the interaction with MAGED1, which
CC participates in the induction of apoptosis (By similarity).
CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues.
CC Phosphorylated by PKC in vitro (By similarity).
CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
CC cleavage does not take place when the receptor is associated with
CC netrin ligand. Its cleavage by caspases is required to induce
CC apoptosis (By similarity).
CC -!- MISCELLANEOUS: Down-regulated in multiple cancers including
CC colorectal, breast, ovary, uterus, stomach, lung, or kidney
CC cancers.
CC -!- SIMILARITY: Belongs to the UNC-5 family.
CC -!- SIMILARITY: Contains 1 death domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 ZU5 domain.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to the
CC presence of introns.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AK131380; BAD18531.1; -.
DR EMBL; BC009333; AAH09333.2; -.
DR EMBL; BC033727; -; NOT_ANNOTATED_CDS.
DR EMBL; AB075856; BAB85562.1; ALT_SEQ.
DR Genew; HGNC:12567; UNC5A.
DR MIM; 607869; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; TSP_1; 1.

DR SMART; SM00409; IG; 1.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 1.
 KW Alternative splicing; Apoptosis; Developmental protein;
 KW Immunoglobulin domain; Phosphorylation; Receptor; Signal;
 KW Transmembrane.
 FT SIGNAL 1 25 Potential.
 FT CHAIN 26 842 Netrin receptor UNC5A.
 FT DOMAIN 1 306 Extracellular (Potential).
 FT TRANSMEM 307 327 Potential.
 FT DOMAIN 328 842 Cytoplasmic (Potential).
 FT DOMAIN 44 141 Ig-like.
 FT DOMAIN 155 234 Ig-like C2-type.
 FT DOMAIN 242 294 TSP type-1.
 FT DOMAIN 439 542 ZU5.
 FT DOMAIN 761 841 Death.
 FT SITE 340 341 Cleavage (by caspase-3) (By similarity).
 FT SITE 605 623 Interaction with DCC (By similarity).
 FT DISULFID 65 124 By similarity.
 FT DISULFID 170 221 By similarity.
 FT CARBOHYD 107 107 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 218 218 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 287 287 N-linked (GlcNAc. . .) (Potential).
 FT VARSPLIC 1 97 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPD
 FT LLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWV
 FT RQVDHVIERSTDGGSN -> MAGTSERSLISSISQPKAIECF
 FT EVKKKAFLTHGRYHGSGATPPKTKDKPKETFCGQT (in
 FT isoform 3).
 FT /FTId=VSP_011693.
 FT VARSPLIC 296 301 TASGPE -> SESSLP (in isoform 2).
 FT /FTId=VSP_011694.
 FT VARSPLIC 302 842 Missing (in isoform 2).
 FT /FTId=VSP_011695.
 SQ SEQUENCE 842 AA; 92958 MW; 3DFADCF973131849 CRC64;

Query Match 92.0%; Score 4410; DB 1; Length 842;
 Best Local Similarity 93.4%; Pred. No. 0;
 Matches 839; Conservative 2; Mismatches 1; Indels 56; Gaps 1;

Qy 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
 Qy 61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGL 120
 ||||||||||||||||||||||||||||:||||||||||||||||||
 Db 61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSNGLPTMEVRINVSQQVEKVFGL 120
 Qy 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIIPPAE 180
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPEGIIPPAE 180
 Qy 181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY 240
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY 240

Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Db	241	-----VDGS	244
Qy	301	WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVA	360
		:	
Db	245	WSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDL DTRNCTSDLCVHTASGPEDVA	304
Qy	361	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHELL	420
Db	305	LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHELL	364
Qy	421	TIQPDLS TTTT TYQGS LCP RQD G P S P K F Q L T N G H L L S P L G G R H T L H H S S P T S E A E E F V S	480
Db	365	TIQPDLS TTTT TYQGS LCP RQD G P S P K F Q L T N G H L L S P L G G R H T L H H S S P T S E A E E F V S	424
Qy	481	RLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHK	540
Db	425	RLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHK	484
Qy	541	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSW	600
Db	485	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSW	544
Qy	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLL FAPVACT	660
Db	545	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLL FAPVACT	604
Qy	661	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW	720
Db	605	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW	664
Qy	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINF	780
Db	665	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINF	724
Qy	781	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL	840
Db	725	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL	784
Qy	841	HLD SHLSFFASKPSP TAMI LNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C	898
Db	785	HLD SHLSFFASKPSP TAMI LNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C	842

RESULT..4.

UN5C_MOUSE

ID UN5C_MOUSE STANDARD; PRT; 931 AA.
 AC O08747; Q8CD16;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Netrin receptor UNC5C precursor (Unc-5 homolog C) (Unc-5 homolog 3)
 DE (Rostral cerebellar malformation protein).
 GN Name=Unc5c; Synonyms=Rcm, Unc5h3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, DISEASE, AND TISSUE
 RP SPECIFICITY.
 RC STRAIN=C57B6/SJL;
 RX MEDLINE=97271898; PubMed=9126743;
 RA Ackerman S.L., Kozak L.P., Przyborski S.A., Rund L.A., Boyer B.B.,
 RA Knowles B.B.;
 RT "The mouse rostral cerebellar malformation gene encodes an UNC-5-like
 RT protein.";
 RL Nature 386:838-842(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sadelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [3]
 RP FUNCTION, AND TISSUE SPECIFICITY.
 RX PubMed=9389662;
 RA Przyborski S.A., Knowles B.B., Ackerman S.L.;
 RT "Embryonic phenotype of Unc5h3 mutant mice suggests chemorepulsion
 RT during the formation of the rostral cerebellar boundary.";
 RL Development 125:41-50(1998).
 RN [4]
 RP INTERACTION WITH DCC.
 RX PubMed=10399920;
 RA Hong K., Hinck L., Nishiyama M., Poo M.-M., Tessier-Lavigne M.,
 RA Stein E.;

RT "A ligand-gated association between cytoplasmic domains of UNC5 and
RT DCC family receptors converts netrin-induced growth cone attraction to
RT repulsion.";
RL Cell 97:927-941(1999).
RN [5]
RP PHOSPHORYLATION SITE TYR-568, AND MUTAGENESIS OF TYR-568.
RX PubMed=11533026; DOI=10.1074/jbc.M103872200;
RA Tong J., Killeen M., Steven R., Binns K.L., Culotti J., Pawson T.;
RT "Netrin stimulates tyrosine phosphorylation of the UNC-5 family of
RT netrin receptors and induces Shp2 binding to the RCM cytodomain.";
RL J. Biol. Chem. 276:40917-40925(2001).
RN [6]
RP FUNCTION.
RX PubMed=12451134; DOI=22/23/10346;
RA Finger J.H., Bronson R.T., Harris B., Johnson K., Przyborski S.A.,
RA Ackerman S.L.;
RT "The netrin 1 receptors Unc5h3 and Dcc are necessary at multiple
RT choice points for the guidance of corticospinal tract axons.";
RL J. Neurosci. 22:10346-10356(2002).
CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates
CC axon repulsion of neuronal growth cones in the developing nervous
CC system upon ligand binding. Axon repulsion in growth cones may be
CC caused by its association with DCC that may trigger signaling for
CC repulsion. Also involved in corticospinal tract axon guidances
CC independently of DCC. It also acts as a dependence receptor
CC required for apoptosis induction when not associated with netrin
CC ligand.
CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O08747-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O08747-2; Sequence=VSP_011702;
CC -!- TISSUE SPECIFICITY: Mainly expressed in regions of differentiating
CC neurons. Highly expressed in brain and lung. Weakly expressed in
CC testis, ovary, spleen, thymus and bladder. Expressed at very low
CC level in kidney, intestine and salivary gland.
CC -!- PTM: Phosphorylated on different cytoplasmic tyrosine residues.
CC Phosphorylation of Tyr-568 leads to an interaction with PTPN11
CC phosphatase, suggesting that its activity is regulated by
CC phosphorylation/dephosphorylation. Tyrosine phosphorylation is
CC netrin-dependent.
CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
CC cleavage does not take place when the receptor is associated with
CC netrin ligand. Its cleavage by caspases is required to induce
CC apoptosis (By similarity).
CC -!- DISEASE: Defects in Unc5c are the cause of rostral cerebellar
CC malformation (Rcm). Rcm is characterized by cerebellar and
CC midbrain defects, apparently as a result of abnormal neuronal
CC migration.
CC -!- SIMILARITY: Belongs to the UNC-5 family.
CC -!- SIMILARITY: Contains 1 death domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.

```

CC      -!- SIMILARITY: Contains 1 ZU5 domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U72634; AAB54103.1; -.
DR      EMBL; AK031655; BAC27495.1; -.
DR      HSSP; P07996; 1LSL.
DR      MGD; MGI:1095412; Unc5c.
DR      GO; GO:0005886; C:plasma membrane; IC.
DR      GO; GO:0005042; F:netrin receptor activity; IDA.
DR      GO; GO:0005515; F:protein binding; IDA.
DR      GO; GO:0007420; P:brain development; IMP.
DR      GO; GO:0030334; P:regulation of cell migration; IMP.
DR      InterPro; IPR000488; Death.
DR      InterPro; IPR011029; DEATH_like.
DR      InterPro; IPR007110; Ig-like.
DR      InterPro; IPR003598; Ig_c2.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR008085; TSP_1.
DR      InterPro; IPR000906; ZU5.
DR      Pfam; PF00531; Death; 1.
DR      Pfam; PF00047; ig; 1.
DR      Pfam; PF00090; TSP_1; 2.
DR      Pfam; PF00791; ZU5; 1.
DR      PRINTS; PR01705; TSP1REPEAT.
DR      SMART; SM00005; DEATH; 1.
DR      SMART; SM00408; IGc2; 1.
DR      SMART; SM00209; TSP1; 2.
DR      SMART; SM00218; ZU5; 1.
DR      PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
DR      PROSITE; PS50835; IG_LIKE; 1.
DR      PROSITE; PS50092; TSP1; 2.
KW      Alternative splicing; Apoptosis; Developmental protein;
KW      Immunoglobulin domain; Phosphorylation; Receptor; Repeat; Signal;
KW      Transmembrane.
FT      SIGNAL          1      40      Potential.
FT      CHAIN           41     931     Netrin receptor UNC5C.
FT      DOMAIN          41     380     Extracellular (Potential).
FT      TRANSMEM        381     401     Potential.
FT      DOMAIN          402     931     Cytoplasmic (Potential).
FT      DOMAIN          62     159     Ig-like.
FT      DOMAIN          161     256     Ig-like C2-type.
FT      DOMAIN          260     314     TSP type-1 1.
FT      DOMAIN          316     368     TSP type-1 2.
FT      DOMAIN          528     631     ZU5.
FT      DOMAIN          850     929     Death.
FT      SITE            415     416     Cleavage (by caspase-3) (By similarity).
FT      SITE            694     712     Interaction with DCC (By similarity).
FT      DISULFID         83     142     By similarity.
FT      DISULFID        188     239     By similarity.
FT      MOD_RES         568     568     Phosphotyrosine.

```

FT	CARBOHYD	236	236	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	361	361	N-linked (GlcNAc. . .) (Potential).
FT	VARSP LIC	370	370	A -> GFIYPISTEHRPQNEYGFSS (in isoform 2).
FT				/FTId=VSP_011702.
FT	MUTAGEN	568	568	Y->F: Abolishes interaction with PTPN11,
FT				leading to a increased level of
FT				phosphorylation.
FT	CONFLICT	16	16	L -> I (in Ref. 2).
FT	CONFLICT	733	733	H -> R (in Ref. 2).
FT	CONFLICT	924	924	S -> Y (in Ref. 2).
SQ	SEQUENCE	931 AA;	103062 MW;	8A5D951A4EECA179 CRC64;

Query Match 58.2%; Score 2787; DB 1; Length 931;
 Best Local Similarity 57.3%; Pred. No. 8.7e-201;
 Matches 522; Conservative 153; Mismatches 208; Indels 28; Gaps 9;

Qy	9	PALLGIVLAAWLRGSGAQQA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC	65
		: : : :	
Db	26	PAL--ALLSASGTGSAAQDDEFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC	83
Qy	66	KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQ	125
		: : : : : :	
Db	84	KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLIVREVSIETSRQVEELFGPEDYWCQ	143
Qy	126	CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLR	185
		: : : : :	
Db	144	CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPGEPVPAEVEWLK	203
Qy	186	NEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVYVNGGW	245
		: : : : : : : :	
Db	204	NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW	263
Qy	246	STWTEWSVCSASCGRGWQKRSRSCCTNPAPLNGGAFCEGQNVQKTACATLCPVDGWSWSPWS	305
		: : : :	
Db	264	STWTEWSVCNSRCGRGYQKRTRCTNPAPLNGGAFCEGQSVQKIACCTTLCVPDGRWTSWS	323
Qy	306	KWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVALYVGL	365
		: : : : : : : :	
Db	324	KWSTCGTECTHWRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPSDDDVALYVGI	383
Qy	366	-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTSQFQPVSIKPSKADNPHLLTIQP	424
		: : : : : : : :	
Db	384	VIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP	440
Qy	425	DLSTTTTTTYQGSCLPRQDGSPKFLTNHLLSPLGGGRHTLHHSS----PTSEAEFEVS	480
		: : : : : : : :	
Db	441	DLTSAAAMYRGFPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSS	499
Qy	481	RLS----TQNYF-----RSLPRGT--SNMTYGTFFNLGGRLMIPNTGISLLIPDAI	526
		: : : : : :	
Db	500	KLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFSNLGGHLIIPNSGVSLIPAGAI	559
Qy	527	PRGKIYEIYLTLHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPSDPS	586
		: : : : : : : : : :	
Db	560	PQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVWSCGPPGALLTRPVILT LHHCADPSTED	619
Qy	587	WSLRLKKQSCGSEWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAA	646

Db 620 WKIQLNQAVQGQWEDVVVVGEENFTTPCYIQLDAAECHILTENLSTYALVGQSTTKAA 679
 Qy 647 KRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH 706
 Db 680 KRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGQQLLEEPKALHFKGSIH 739
 Qy 707 NLRLSIHDVPSSLWWSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWV 766
 Db 740 NLRLSIHDIAHSLWWSKLLAKYQEIPFYHIWSGSQRNHCTFTLERLSLNTVELVCKLCV 799
 Qy 767 WQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPP 826
 Db 800 RQVEGEGQIFQLNCTVSEETPIDPLLDPASTITVTGSAFSIPLPIRQKLCSSLDAP 859
 Qy 827 CRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQP 886
 Db 860 QTRGHDWRMLAHKLNLDRLNYFATKSSPTGVILDLWEAQNFDPGNLSMLAAVLEEMGRH 919
 Qy 887 DAGLFTVSEAE 897
 : : : :
 Db 920 ETVVSLAEGQ 930

RESULT 5

UN5C_RAT

ID UN5C_RAT STANDARD; PRT; 931 AA.
 AC Q761X5;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Netrin receptor UNC5C precursor (Unc-5 homolog C) (Unc-5 homolog 3).
 GN Name=Unc5c; Synonyms=Unc5h3;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND DISEASE.
 RX PubMed=15010202; DOI=10.1016/j.molbrainres.2003.12.003;
 RA Kuramoto T., Kuwamura M., Serikawa T.;
 RT "Rat neurological mutations cerebellar vermis defect and hobble are
 RT caused by mutations in the netrin-1 receptor gene Unc5h3."
 RL Brain Res. Mol. Brain Res. 122:103-108(2004).
 RN [2]
 RP FUNCTION.
 RX PubMed=11387206; DOI=10.1093/emboj/20.11.2715;
 RA Llambe F., Causeret F., Bloch-Gallego E., Mehlen P.;
 RT "Netrin-1 acts as a survival factor via its receptors UNC5H and DCC."
 RL EMBO J. 20:2715-2722(2001).
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates
 CC axon repulsion of neuronal growth cones in the developing nervous
 CC system upon ligand binding. Axon repulsion in growth cones may be
 CC caused by its association with DCC that may trigger signaling for
 CC repulsion. Also involved in corticospinal tract axon guidances
 CC independently of DCC. It also acts as a dependence receptor
 CC required for apoptosis induction when not associated with netrin

CC ligand.
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: Mainly expressed in brain. Also expressed in
 CC kidney. Not expressed in developing or adult lung.
 CC -!- PTM: Phosphorylated on different cytoplasmic tyrosine residues.
 CC Phosphorylation of Tyr-568 leads to an interaction with PTPN11
 CC phosphatase, suggesting that its activity is regulated by
 CC phosphorylation/dephosphorylation. Tyrosine phosphorylation is
 CC netrin-dependent (By similarity).
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
 CC cleavage does not take place when the receptor is associated with
 CC netrin ligand. Its cleavage by caspases is required to induce
 CC apoptosis.
 CC -!- DISEASE: Defects in Unc5c are the cause of cerebellar vermis
 CC defect (cvd) and hobble (hob) phenotypes. Cvd and hob rats exhibit
 CC cerebellar and midbrain defects, possibly as a result of abnormal
 CC neuronal migration, and exhibit laminar structure abnormalities in
 CC the fused cerebellar hemispheres and ectopic cerebellar tissues in
 CC the cerebello-pontine junction.
 CC -!- SIMILARITY: Belongs to the UNC-5 family.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 ZU5 domain.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB118026; BAD05181.1; -.
 DR RGD; 735109; Unc5c.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR011029; DEATH_like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; Death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; TSP_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.

DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Apoptosis; Developmental protein; Immunoglobulin domain;
 KW Phosphorylation; Receptor; Repeat; Signal; Transmembrane.
 FT SIGNAL 1 40 Potential.
 FT CHAIN 41 931 Netrin receptor UNC5C.
 FT DOMAIN 41 380 Extracellular (Potential).
 FT TRANSMEM 381 401 Potential.
 FT DOMAIN 402 931 Cytoplasmic (Potential).
 FT DOMAIN 62 159 Ig-like.
 FT DOMAIN 161 256 Ig-like C2-type.
 FT DOMAIN 260 314 TSP type-1 1.
 FT DOMAIN 316 368 TSP type-1 2.
 FT DOMAIN 528 631 ZU5.
 FT DOMAIN 850 929 Death.
 FT SITE 415 416 Cleavage (by caspase-3) (By similarity).
 FT SITE 694 712 Interaction with DCC (By similarity).
 FT DISULFID 83 142 By similarity.
 FT DISULFID 188 239 By similarity.
 FT MOD_RES 568 568 Phosphotyrosine (By similarity).
 FT CARBOHYD 236 236 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 361 361 N-linked (GlcNAc. . .) (Potential).
 SQ SEQUENCE 931 AA; 103134 MW; 25B183A97BCB8401 CRC64;

Query Match 57.8%; Score 2770; DB 1; Length 931;
 Best Local Similarity 57.0%; Pred. No. 1.7e-199;
 Matches 519; Conservative 154; Mismatches 210; Indels 28; Gaps 9;

Qy 9 PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
 ||| :|| || | : | : ||||:|:| : ||||| | |
 Db 26 PAL--ALLSASGTGSAAQDDDFHHELPETFPSPPEPLPHFLIEPEEAYIVKNKPVNLYC 83
 Qy 66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQ 125
 || ||||:|||| || | |||: | :|| | | :|||:|:| | :|||
 Db 84 KASPATQIYFKCNSEWVHQKDHVDERVDETSGLIVREVSIIEISRQVEELFGPEDYWCQ 143
 Qy 126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLR 185
 |||||:||||:|:|:| || | |||| | ||||| :| | ||||| | ||||:|
 Db 144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPGEGIPMAEVEWLK 203
 Qy 186 NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
 |||:| | | | :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:| :| | |||||
 Db 204 NEDIIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263
 Qy 246 STWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGWSWPWS 305
 ||| ||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:| | | ||||| | :| |
 Db 264 STWAEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACCTLCPVDGRWTSWS 323
 Qy 306 KWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGL 365
 ||| | :||| | | : |||:|:|:| | | :||| | :| :| |||||:|
 Db 324 KWSTCGTECTHWRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDVALYVGI 383
 Qy 366 -IAVAVCLVLLLVLILVYCRKKEGLSDVDASSILTSGFQPVSIKPSKADNPHELLTIQP 424
 ||| ||| : :| : || : :|:| || | ||||:| :| | | :|
 Db 384 VIAVTVCLAITVVVALFVYRKNHRDFESNIIDSSALNGGFQPVNIKAARQD---LLAVPP 440
 Qy 425 DLSTTTTTYQGSILCPQDGPSPKFQLTNGHLLSPLGGGRHTLHHSS----PTSEAEFVS 480

Db	441	DLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSS	499
Qy	481	RLS---TQNYF-----RSLPRGT--SNMTYGTFFNLGGRLMIPNTGISLLIPPDAI	526
Db	500	KLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFFNSLGGHLIIPNSGVSLIPAGAI	559
Qy	527	PRGKIYEIYLTTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDPS	586
Db	560	PQGRVYEMYVTVHRKENMRPPMEDSQTLTPVWSCGPPGALLTRPVILTLLHHCADPNTE	619
Qy	587	WSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA	646
Db	620	WKIQLKNQAVQGQWEDVVVVGEEFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAA	679
Qy	647	KRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH	706
Db	680	KRLKLAIFGPLCCSSLEYSIRVYCLDDTDALKEVLQLERQMGGQLLEPKALHFKGSIH	739
Qy	707	NLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWV	766
Db	740	NLRLSIHDITHSLWKSLLAKYQEIPFYHIWSGSQRNHCTFTLERLSLNTVELVCKLCV	799
Qy	767	WQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKISSLDP	826
Db	800	RQVEGEGQIFQLNCTVSEPTGIDLPLLDPASTITTTVTGPSAFSIPPIRQKLCSSLDAP	859
Qy	827	CRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQP	886
Db	860	QTRGHDWRMLAHKLNLDRLNYFATKSSPTGVILDLWEAQNFDPGNLSMLAAVLEEMGRH	919
Qy	887	DAGLFTVSEAE	897
Db	920	ETVVSLLAEGQ	930

RESULT 6

UN5C_CHICK

ID UN5C_CHICK STANDARD; PRT; 931 AA.

AC Q7T2Z5;

DT 25-OCT-2004 (Rel. 45, Created)

DT 25-OCT-2004 (Rel. 45, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Netrin receptor UNC5C precursor (Unc-5 homolog C) (Unc-5 homolog 3)

DE (cUNC-5H3).

GN Name=UNC5C;

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

RX PubMed=12799087;

RA Guan W., Condic M.L.;

RT "Characterization of Netrin-1, Neogenin and cUNC-5H3 expression during chick dorsal root ganglia development.";

RL Gene Expr. Patterns 3:369-373(2003).
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates
 CC axon repulsion of neuronal growth cones in the developing nervous
 CC system upon ligand binding (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: Restricted to proprioceptive neurons.
 CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the UNC-5 family.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 ZU5 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AY187310; AAO67275.1; -.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; Death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; TSP_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Developmental protein; Immunoglobulin domain; Phosphorylation;
 KW Receptor; Repeat; Signal; Transmembrane.
 FT SIGNAL 1 39 Potential.
 FT CHAIN 40 931 Netrin receptor UNC5C.
 FT DOMAIN 40 380 Extracellular (Potential).
 FT TRANSMEM 381 401 Potential.
 FT DOMAIN 402 931 Cytoplasmic (Potential).
 FT DOMAIN 62 159 Ig-like.
 FT DOMAIN 161 256 Ig-like C2-type.
 FT DOMAIN 260 314 TSP type-1 1.
 FT DOMAIN 316 368 TSP type-1 2.
 FT DOMAIN 528 631 ZU5.
 FT DOMAIN 850 929 Death.
 FT DISULFID 83 142 By similarity.
 FT DISULFID 188 239 By similarity.
 FT CARBOHYD 236 236 N-linked (GlcNAc . .) (Potential).
 FT CARBOHYD 361 361 N-linked (GlcNAc . .) (Potential).

SQ SEQUENCE 931 AA; 102906 MW; 1E23A0D84F2E2C62 CRC64;

Query Match 57.6%; Score 2761; DB 1; Length 931;
Best Local Similarity 57.0%; Pred. No. 7.9e-199;
Matches 518; Conservative 151; Mismatches 212; Indels 28; Gaps 9;

```
Qy      9 PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
      ||| ||| ||| ||| : | |: ||||:||||: ||||| |||
Db     26 PAL--AVLGASRPGSAAQDDDDFFHELPETFPSPDPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy     66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQ 125
      || ||||:|||| ||| | |||: | :|| | | :|||:|:| | :|||
Db     84 KASPATQIYFKCNSEWVHQKDHVVDERVDETSGLVCEVSIEISRQVEELFGPEDYWCQ 143

Qy    126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEWLR 185
      |||||:||||:|||| ||| ||||| ||||| :| ||||| |||||:
Db    144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPGIPVAEVEWLK 203

Qy    186 NEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW 245
      ||::|| | | || :|:|::|||:||||| |||||:|:| :| |||||
Db    204 NEEVIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy    246 STWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGWSWSPWS 305
      ||||| |:: |||:|||:|:||||| ||||| || ||||| |: ||
Db    264 STWTEWSACNSRCGRGFQKRTRCTNPAPLNGGAFCEGQNVQKIACTTLCPPVDGKWTWS 323

Qy    306 KWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGL 365
      ||| || :||| |||: |||:|||:|:| | :||| ||: :| :|||:|:
Db    324 KWSTCGTECTHWRRRECTAPAPKNGGKDCEGLVLQSKNCTDGLCMQAAPSDDDVALYVGI 383

Qy    366 -IAVAVCLVLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHELLTIQP 424
      ||| ||| : :| : || : ||: ||| | ||||:| :| | :|
Db    384 VIAVIVCLAISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP 440

Qy    425 DLSTTTTTTYQGSCLPRQDGSPKPFQLTNGHLLSPLGGGRHTLHHSS----PTSEAEFVS 480
      ||:: ||: | : | | :|| :| || : :::| | | :| |
Db    441 DLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNTSGAVTPQDELSDFSS 499

Qy    481 RLS---TQNYF-----RSLPRGT--SNMTYGTNFNLGGRLMIPNTGISLLIPPDAI 526
      :|| ||: :|| | | :||| ||| |:|||:|:|||| |:
Db    500 KLSPQITQSLLENETLNVKNQSLARQTDPSCTAFGTFNLSLGGHLVIPNSGVSLIPAGAV 559

Qy    527 PRGKIYEIYLTTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDPS 586
      |:|:|:|:|:|:| :| |: |||:|:||||| |||||:| | || |:|
Db    560 PQGRVYEMYVTVHRKEGMRPPVEDSQTLLTPVVS CGPPGALLTRPVVLTMHCAEPNMDD 619

Qy    587 WSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAA 646
      | :|| |: :| ||||: ||| : | ||: ||: || | :|||:|:| |||
Db    620 WQIQLKHQAGQGPWEDVVVGEENFTTPCYIQLDPEACHILTETLSTYALVQSITKAAA 679

Qy    647 KRLKLLLFPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH 706
      |||| | |:|:|:|:|:| || |||||:|:|:|:|:|:|:| ||| | |
Db    680 KRLKLAIFGPLSCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEPKTLHFKGSTH 739

Qy    707 NLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWV 766
      |||||: ||||| |||||:| || ||||| | :| :| ||| |
Db    740 NLRLSIHDIAHSLWKSLLPAKYQEIPFYHIWSGQRNLHCTFTLERFSLNLTLELVCKLCV 799
```

Qy 767 WQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPP 826
 |||:| | :| :::: : ::| : :||:| | | |||: ||| |
 Db 800 RQVEGEGQIFQLNCVSEETPIDYPMDSAGSITTIVGPNAFSIPLPIRQKLCSSLDAP 859
 Qy 827 CRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQP 886
 || ||| || || || :|:|:|:| ||| :|:|:|:|:|:|:| ||| : :|:
 Db 860 QTRGHDWRMLAHKLKLDRLNYFATKSSPTGVILDLWEAQNFDPGNLSMLAAVLEEMGRH 919
 Qy 887 DAGLFTVSE 895
 : : :|
 Db 920 ETVVSLAAE 928

RESULT 7

UN5C_HUMAN

ID UN5C_HUMAN STANDARD; PRT; 931 AA.
 AC O95185; Q8IUT0;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Netrin receptor UNC5C precursor (Unc-5 homolog C) (Unc-5 homolog 3).
 GN Name=UNC5C; Synonyms=UNC5H3;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
 RC TISSUE=Brain;
 RX MEDLINE=99000841; PubMed=9782087; DOI=10.1006/geno.1998.5425;
 RA Ackerman S.L., Knowles B.B.;
 RT "Cloning and mapping of the UNC5C gene to human chromosome 4q21-q23.";
 RL Genomics 52:205-208(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Lung;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP DOWN-REGULATION IN CANCER.
 RX PubMed=12655055; DOI=10.1073/pnas.0738063100;
 RA Thiebault K., Mazelin L., Pays L., Llambi F., Joly M.-O.,
 RA Scoazec J.-Y., Saurin J.-C., Romeo G., Mehlen P.;
 RT "The netrin-1 receptors UNC5H are putative tumor suppressors
 RT controlling cell death commitment.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:4173-4178(2003).
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates
 CC axon repulsion of neuronal growth cones in the developing nervous
 CC system upon ligand binding. Axon repulsion in growth cones may be
 CC caused by its association with DCC that may trigger signaling for
 CC repulsion. Also involved in corticospinal tract axon guidances
 CC independently of DCC. It also acts as a dependence receptor
 CC required for apoptosis induction when not associated with netrin
 CC ligand (By similarity).
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=O95185-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O95185-2; Sequence=VSP_011700, VSP_011701;
 CC -!- TISSUE SPECIFICITY: Mainly expressed in brain. Also expressed in
 CC kidney. Not expressed in developing or adult lung.
 CC -!- PTM: Phosphorylated on different cytoplasmic tyrosine residues.
 CC Phosphorylation of Tyr-568 leads to an interaction with PTPN11
 CC phosphatase, suggesting that its activity is regulated by
 CC phosphorylation/dephosphorylation. Tyrosine phosphorylation is
 CC netrin-dependent (By similarity).
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
 CC cleavage does not take place when the receptor is associated with
 CC netrin ligand. Its cleavage by caspases is required to induce
 CC apoptosis (By similarity).
 CC -!- MISCELLANEOUS: Down-regulated in multiple cancers including
 CC colorectal, breast, ovary, uterus, stomach, lung, or kidney
 CC cancers.
 CC -!- SIMILARITY: Belongs to the UNC-5 family.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 ZU5 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF055634; AAC67491.1; -.
 DR EMBL; BC041156; AAH41156.1; -.

DR HSSP; P07996; 1LSL.
 DR Genew; HGNC:12569; UNC5C.
 DR MIM; 603610; -.
 DR GO; GO:0005042; F:netrin receptor activity; TAS.
 DR GO; GO:0007411; P:axon guidance; TAS.
 DR GO; GO:0007420; P:brain development; TAS.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR011029; DEATH_like.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; Death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; TSP_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Alternative splicing; Apoptosis; Developmental protein;
 KW Immunoglobulin domain; Phosphorylation; Polymorphism; Receptor;
 KW Repeat; Signal; Transmembrane.
 FT SIGNAL 1 40 Potential.
 FT CHAIN 41 931 Netrin receptor UNC5C.
 FT DOMAIN 41 380 Extracellular (Potential).
 FT TRANSMEM 381 401 Potential.
 FT DOMAIN 402 931 Cytoplasmic (Potential).
 FT DOMAIN 62 159 Ig-like.
 FT DOMAIN 161 256 Ig-like C2-type.
 FT DOMAIN 260 314 TSP type-1 1.
 FT DOMAIN 316 368 TSP type-1 2.
 FT DOMAIN 528 631 ZU5.
 FT DOMAIN 850 929 Death.
 FT SITE 415 416 Cleavage (by caspase-3) (By similarity).
 FT SITE 694 712 Interaction with DCC (By similarity).
 FT DISULFID 83 142 By similarity.
 FT DISULFID 188 239 By similarity.
 FT MOD_RES 568 568 Phosphotyrosine (By similarity).
 FT CARBOHYD 236 236 N-linked (GlcNAc . . .) (Potential).
 FT CARBOHYD 361 361 N-linked (GlcNAc . . .) (Potential).
 FT VARSPLIC 370 370 T -> SFIYPISTEQRTQNEYGFSS (in isoform 2).
 FT /FTId=VSP_011700.
 FT VARSPLIC 579 931 Missing (in isoform 2).
 FT /FTId=VSP_011701.
 FT VARIANT 37 37 G -> V (in dbSNP:2306715).
 FT /FTId=VAR_019731.
 FT VARIANT 721 721 T -> M (in dbSNP:2289043).
 FT /FTId=VAR_019732.
 FT CONFLICT 219 219 T -> I (in Ref. 1).
 FT CONFLICT 489 489 S -> T (in Ref. 1).
 SQ SEQUENCE 931 AA; 103101 MW; EFD71122C98DABB8 CRC64;

Query Match 57.5%; Score 2755; DB 1; Length 931;
Best Local Similarity 56.4%; Pred. No. 2.2e-198;
Matches 514; Conservative 154; Mismatches 215; Indels 28; Gaps 9;

Qy	9	PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC	65
		: : : : :	
Db	26	PAL--ALLSASGTGSAAQDDDDFFHELPETFPSPDPEPLPHFLIEPEEAYIVKNKPVNLYC	83
Qy	66	KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQ	125
		: :: : : : : :	
Db	84	KASPATQIYFKCNSEWVHQKDHIVDERVDETSGLIVREVSIEISRQVEELFGPEDYWCQ	143
Qy	126	CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLR	185
		: : : : :	
Db	144	CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK	203
Qy	186	NEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGW	245
		:: : :: : : : : : :	
Db	204	NEDIIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW	263
Qy	246	STWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWS	305
		:: : : : : : : :	
Db	264	STWTEWSVCNSRCGRGYQKRTRCTNPAPLNGGAFCEGQSVQKIACTTLCVPDGRWTPWS	323
Qy	306	KWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVALYVGL	365
		: : : : : : : : :	
Db	324	KWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQTAPDSDDVALYVGI	383
Qy	366	-IAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHELLTIQP	424
		: : : : : : : :	
Db	384	VIAVIVCLAISVVVALFVYRKNNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP	440
Qy	425	DLSTTTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHS----SPTSEAEFVS	480
		:: : : : : : : :	
Db	441	DLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNTSGAVSPQDDLSEFTS	499
Qy	481	RLS---TQNYF-----RSLPRGT--SNMTYGTFFNLGGRLMIPNTGISLLIPPDAI	526
		: : : : :	
Db	500	KLSPQMTQSLENEALS LKNQSLARQTDPSCTAFGSFNSLGGHLIVPNSGVSLIPAGAI	559
Qy	527	PRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDPS	586
		: :: : : : : : : : : : :	
Db	560	PQGRVYEMYVTVHRKETMRPPMDDSQTLLTPVWSCGPPGALLTRPVVLTMHHCADPNTE	619
Qy	587	WSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAA	646
		: : : : : : : : : :	
Db	620	WKILLKNQAAQGQWEDVVVGEENFTTPCYIKLDAEACHILTENLSTYALVGHSTTKAAA	679
Qy	647	KRLKLLL FAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH	706
		: : : : : : :	
Db	680	KRLKLAIFGPLCCSSLEYSIRVYCLDDTDALKEILHLERQTGGQLLEPKALHFKGSTH	739
Qy	707	NLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWV	766
		: : : : : :	
Db	740	NLRLSIHDIAHSLWKS KLLAKYQEIPFYHVWSGSQRNLHCTFTLERFSLNTVELVCKLCV	799

Qy 767 WQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPP 826
 ||||:| | :| :::: :| |: : : |||| | | |||: ||| |
 Db 800 RQVEGEGQIFQLNCTVSEPTGIDLPLDPANTITTTVTGPSAFSIPLPIRQKLCSSLDAP 859
 Qy 827 CRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQP 886
 || ||| || ||:| | :|::|:| ||| :||:||||:|:|:| ||| : :|:
 Db 860 QTRGHDWRMLAHKLNLDRLNYFATKSSPTGVILDLWEAQNFDPGNLSMLAAVLEEMGRH 919
 Qy 887 DAGLFTVSEAE 897
 : : :| :
 Db 920 ETVVSLAAEQ 930

RESULT 8

UN5B_XENLA

ID UN5B_XENLA STANDARD; PRT; 943 AA.
 AC Q8JGT4;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Netrin receptor UNC5B precursor (UNC-5 homolog) (Protein XUNC-5).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=22239703; PubMed=12351179; DOI=10.1016/S0925-4773(02)00215-0;
 RA Anderson R.B., Holt C.E.;
 RT "Expression of UNC-5 in the developing Xenopus visual system.";
 RL Mech. Dev. 118:157-160(2002).
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates
 CC axon repulsion of neuronal growth cones in the developing nervous
 CC system upon ligand binding (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: In the developing visual system, it is
 CC expressed within the developing optic vesicles and later become
 CC restricted to the dorsal ciliary marginal zone, a site of
 CC retinoblast proliferation and differentiation.
 CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the UNC-5 family.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 ZU5 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AY099459; AAM34486.1; -.
 DR HSSP; P07996; 1LSL.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR011029; DEATH_like.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; Death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; TSP_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Developmental protein; Immunoglobulin domain; Phosphorylation;
 KW Receptor; Repeat; Signal; Transmembrane.
 FT SIGNAL 1 30 Potential.
 FT CHAIN 31 943 Netrin receptor UNC5B.
 FT DOMAIN 31 380 Extracellular (Potential).
 FT TRANSMEM 381 401 Potential.
 FT DOMAIN 402 943 Cytoplasmic (Potential).
 FT DOMAIN 51 148 Ig-like.
 FT DOMAIN 150 245 Ig-like C2-type.
 FT DOMAIN 249 303 TSP type-1 1.
 FT DOMAIN 305 357 TSP type-1 2.
 FT DOMAIN 540 643 ZU5.
 FT DOMAIN 863 941 Death.
 FT DISULFID 72 131 By similarity.
 FT DISULFID 177 228 By similarity.
 FT CARBOHYD 225 225 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 350 350 N-linked (GlcNAc. . .) (Potential).
 SQ SEQUENCE 943 AA; 105083 MW; A024E24A7EDB6175 CRC64;

Query Match 55.2%; Score 2646.5; DB 1; Length 943;
 Best Local Similarity 53.0%; Pred. No. 3.4e-190;
 Matches 496; Conservative 163; Mismatches 229; Indels 47; Gaps 8;

Qy 10 ALLGIVLAAWLRG-----SGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVL 62
 :| | : :| : | : : | | : |||||:|||| :|||
 Db 10 AALAAILVALILSCNFPSSSTAGIEYSDVLPDSFPSAPAESLPHFLLEPEDAYIVKNKPVE 69
 Qy 63 LVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGLLEEEY 122
 ||||| :|||:||||||| | ||: : | :|| ||:| |||||:|:||||:|
 Db 70 LVCKANPATQIYFKCNGEWVNQNDHITKERVDDVTGLVVREVQIEVSRQQVEELFGLEDY 129
 Qy 123 WCQCVAWSSSGTTKSKQAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVE 182
 |||||:||||:|:|:| ||||:|||| ||:| | :| |||||:|||||
 Db 130 WCQCVAWSSAGTTKSKRSYVRIAYLRKNFDQEPLGKEVALEQEALLQCRPPEGVPPAEVE 189
 Qy 183 WLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVN 242

Db	190	WLKNEEIIDPTKDTNFLTIDHNLIIKQARLSDTANYTCVSKNIVAKRRSTTATVIVFVN	249
Qy	243	GGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWS	302
Db	250	GGWSSWTEWSPCNNRCGHGWQKRTRCTNPAPLNGGTMCEGQQYQKFACNTMCPVDGGWT	309
Qy	303	PWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHS-----	352
Db	310	EWKWSACSTECHWSRECNAPTCKNGGKDCSGMLLDSKNCTDGLCMQNKRVLGETKSR	369
Qy	353	-ASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSS-ILTSGFQPVSI	409
Db	370	LLESTGDVALYAGLVVAIFIVIILLMAVGIVVYRRNCREFDTDITDSSAALTGGFHPVNF	429
Qy	410	KPSKADNPHLL--TIQPDLSSTTTTYQGSCLPRQDGSPKQFQLTNGHLLSPLGGGRHTLH	467
Db	430	KTSRHDNSQLIHPAMQPDLTANAGIYRGNMYALQDS-ADKIPMTNSPLLDPLPNLKIKVY	488
Qy	468	HSS-----PTSEAEFVSRLSTQN-----YFRSLPRGTSNMTYGTGTF	503
Db	489	NSSTVGSSPGIHDGNNLLGKTGTPTSDNNIMNARNKNMSMQHLLTLPRDSSNSVTGTGTF	548
Qy	504	NFLGGRMLIPNTGISLLIPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGP	563
Db	549	GSLGGRLTFPNTGVSLLIPOGAIPQKGYYEYMLINKRENTVLPSEGTQTILSPIITCGP	608
Qy	564	PGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASA	623
Db	609	TGLLLCKPVILTVPHCADINTSDWILQLKTQSHQGNWEEVVTLNEETLNTPCYCQLES	668
Qy	624	CYVFTEQLGRFALVGEALSVAALKRLKLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	683
Db	669	CHTLDDQLGTYAFVGESYSRSAIKRLQLAIFAPMLCTSLEYNLKVYCVEDTPDALKEVLE	728
Qy	684	LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRY	743
Db	729	LEKTLGGYLVEEPKLLMFKDSYHNLRLSIHDIPHSLWRSKLMAKYQEIPFYHIWSGSQRT	788
Qy	744	LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	803
Db	789	LHCTFTLERYSLAATELTCKICVRQVEGEGQIFQLHTLLEENVKSFDPFCSQAENSVTTH	848
Qy	804	VGPSAFKIPFLIRQKIISLDPCCRRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLW	863
Db	849	LGPYAFKIPFSIRQKICNSLDAPNSRGNDWRLLAQKLCMDRYLNYFATKASPTGVILDW	908
Qy	864	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	898
Db	909	EALHQDDGDLNLTLASALEEMGKSEMMLVMATDGD	943

RESULT 9

UN5B_MOUSE

ID UN5B_MOUSE STANDARD; PRT; 945 AA.
AC Q8K1S3; Q6PFH0; Q80Y85; Q9D398;
DT 25-OCT-2004 (Rel. 45, Created)

DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Netrin receptor UNC5B precursor (Unc-5 homolog B) (Unc-5 homolog 2).
 GN Name=Unc5b; Synonyms=Unc5h2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
 RX MEDLINE=22239710; PubMed=12351186; DOI=10.1016/S0925-4773(02)00248-4;
 RA Engelkamp D.;
 RT "Cloning of three mouse unc-5 genes and their expression patterns at
 RT mid-gestation.";
 RL Mech. Dev. 118:191-197(2002).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP TISSUE SPECIFICITY.
 RX PubMed=12799072;
 RA Dalvin S., Anselmo M.A., Prodhan P., Komatsuzaki K., Schnitzer J.J.,
 RA Kinane T.B.;
 RT "Expression of Netrin-1 and its two receptors DCC and UNC5H2 in the
 RT developing mouse lung.";
 RL Gene Expr. Patterns 3:279-283(2003).
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates
 CC axon repulsion of neuronal growth cones in the developing nervous
 CC system upon ligand binding. Axon repulsion in growth cones may be
 CC caused by its association with DCC that may trigger signaling for
 CC repulsion. It also acts as a dependence receptor required for
 CC apoptosis induction when not associated with netrin ligand (By
 CC similarity).
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC. Interacts
 CC with GNAI2 via its cytoplasmic part (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8K1S3-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8K1S3-2; Sequence=VSP_011699;
 CC -!- TISSUE SPECIFICITY: Highly expressed in brain. Expressed in lung
 CC during late development. Expressed during early blood vessel
 CC formation, in the semicircular canal and in a dorsal to ventral
 CC gradient in the retina.
 CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By
 CC similarity).
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
 CC cleavage does not take place when the receptor is associated with
 CC netrin ligand. Its cleavage by caspases is required to induce
 CC apoptosis (By similarity).
 CC -!- SIMILARITY: Belongs to the UNC-5 family.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 ZU5 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

DR EMBL; AJ487853; CAD32251.1; -.
DR EMBL; AK018177; BAB31108.1; -.
DR EMBL; BC048162; AAH48162.1; ALT_INIT.
DR EMBL; BC057560; AAH57560.1; -.
DR HSSP; P07996; 1LSL.
DR MGD; MGI:894703; Unc5b.
DR InterPro; IPR000488; Death.
DR InterPro; IPR011029; DEATH_like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00531; Death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50092; TSP1; 2.
KW Alternative splicing; Apoptosis; Developmental protein;
KW Immunoglobulin domain; Phosphorylation; Receptor; Repeat; Signal;
KW Transmembrane.
FT SIGNAL      1      26      Potential.
FT CHAIN       27     945     Netrin receptor UNC5B.
FT DOMAIN      27     377     Extracellular (Potential).
FT TRANSMEM    378     398     Potential.
FT DOMAIN      399     945     Cytoplasmic (Potential).
FT DOMAIN      48     145     Ig-like.
FT DOMAIN     153     242     Ig-like C2-type.
FT DOMAIN     246     300     TSP type-1 1.
FT DOMAIN     302     354     TSP type-1 2.
FT DOMAIN     541     644     ZU5.
FT DOMAIN     865     943     Death.
FT SITE        412     413     Cleavage (by caspase-3) (By similarity).
FT SITE        707     725     Interaction with DCC (By similarity).
FT DISULFID     69     128     By similarity.
FT DISULFID    174     225     By similarity.
FT CARBOHYD    222     222     N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD    347     347     N-linked (GlcNAc. . .) (Potential).
FT VARSPLIC    356     367     NQRTLNDPKSHP -> T (in isoform 2).
FT                                     /FTId=VSP_011699.
FT CONFLICT    238     238     T -> A (in Ref. 2).
FT CONFLICT    394     394     V -> E (in Ref. 2).
FT CONFLICT    679     679     T -> S (in Ref. 2).

```

FT CONFLICT 874 874 N -> D (in Ref. 2).
SQ SEQUENCE 945 AA; 103738 MW; 80E896F0F0E06012 CRC64;

Query Match 53.8%; Score 2578.5; DB 1; Length 945;
Best Local Similarity 53.2%; Pred. No. 4.5e-185;
Matches 506; Conservative 150; Mismatches 235; Indels 61; Gaps 15;

```
Qy      1 MAVRPGLWPALLGIVLAAW-----LRG--SGAQQSATVANPVPGANPDLLPHFLVEPEDV 53
      | | |: ||| :| |   | | | |   : : | | : ||:|:|:|
Db      1 MRARSGVRSALLLALLLCWDPTPSLAGVDSAGQ---VLPDSYPSAPAEQLPYFLLEPQDA 57

Qy     54 YIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQV 113
      ||||| | | |:| ||||:||||| | ||| : | | :|| | ||:| |||||
Db     58 YIVKNKPVLEHCRAPATQIYFKCNGEWVSQNDHVTQESLDEATGLRVREVQIEVSRQQV 117

Qy    114 EKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPP 173
      |:| ||||:||||| ||||:||||| ||||:||||| |: :| |||
Db    118 EELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPP 177

Qy    174 EGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSA 233
      ||:| ||||:||||:|:| | | :| :|:|:| ||||:||||| ||||:|
Db    178 EGVFPAEVEWLKNEDVIDPAQDTNFLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRST 237

Qy    234 SAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPLNGGAFCEGQNVQKTACAT 293
      :| ||||| ||||:| ||| || | |||||:|:| ||||| ||||| |
Db    238 TATVIVYVNGGWSSWAIEWSPCSNRCGRGWQKRTRCTNPAPLNGGAFCEGQAFQKTACTT 297

Qy    294 LCPVDGSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCV--- 350
      :|||:|:| ||||| :| ||||| | |:| | :| || |:| | ||
Db    298 VCPVDGAWTEWSKWSACSTCAHWSRECMAPPQNGGRDCSGTLLDSKNCTDGLCVLNQ 357

Qy    351 -----HSASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLSDVADSS-IL 400
      |   |||| || :| | : :| :| :| | | | | :|:| || |
Db    358 RTLNDPKSHPLETSGDVALYAGLVAVFVVAVLMAVGVIYRRNCRDFDITDSSAAL 417

Qy    401 TSGFQPVSIKPSKADNPHELL--TIQPDLSSTTTTTTYQGSCLPRQDGPSPKFQLTNGHLLSP 458
      | || ||: | :| :| || : |||: : | :| : || : | :| || |
Db    418 TGGFHPVNFKTARPNNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLD 476

Qy    459 LGGGRHTLHHSS-----PTSEAEFVSRLSTQNYFRS----- 490
      | : :|:| | | : :| || : :| |
Db    477 LPSLKIKVYNSSTIGSGSLADGADLLGVLPPTYPGDF-SRDTHFLHLRSASLSQHLL 535

Qy    491 -LPRGTSNMTYGTFFNLGGRMLIPNTGISLLIPDAIPRGKIYEIYLTLHKPEDVRLPLA 549
      ||| |: ||| |||| :| ||:|:| | ||:| | |:| | :| | ||:
Db    536 GLPRDPSSSVSGTFCGLGGRLSLPGTGVSLVPNGAIPQGFYDLYLHINKAEST-LPLS 594

Qy    550 -GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCGWSWEDVLHLGE 608
      | ||:| | |:| | ||:| : || | | :| | :| | ||:| | |
Db    595 EGSQTVLSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIFQLKTQAHQHWEEVVTLDE 654

Qy    609 EAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLL FAPVACTSLEYNIRV 668
      | : ||||| :|:| :||| : :||: | :| |||:| :|| | ||||:|
Db    655 ETLNTPCYCQLEAKSCHILLDQLGTYVFMGESYSRSAVKRLQLAIFAPALCTSLEYSLRV 714

Qy    669 YCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWWSKLLVSY 728
      ||| || | ||||:|:| ||| |:|:| : | ||||| ||||:|:| : |:| || |
```

Db 715 YCLEDETPVALKEVLELERTLGGYLVEEPKPLLFKDSYHNLRLSLHDI PHAHWRSKLLAKY 774

Qy 729 QEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRF 788
 |||||:|:|:| ||||| | ::: | | : | |||:| | : : :|

Db 775 QEIPFYHVWNGSQRALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPA 833

Qy 789 AELLALESEAG--VPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShL 846
 | | | | : : | | | | | | | | | | | | | | : | : |

Db 834 GSLDALCSAPGNAITTLQLGPYAFKIPLSIRQKICSSLDAPNSRGNDWRLLAQKLSMDRYL 893

Qy 847 SFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898
 : : | : | : | : | : | : | : | : | : | : | : |

Db 894 NYFATKASPTGVILDLWEARQQDDGDLNSLASALEEMGKSEMLVAMATDGDC 945

RESULT 10

UN5B_RAT

ID UN5B_RAT STANDARD; PRT; 945 AA.

AC 008722;

DT 25-OCT-2004 (Rel. 45, Created)

DT 25-OCT-2004 (Rel. 45, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Netrin receptor UNC5B precursor (Unc-5 homolog B) (Unc-5 homolog 2).

GN Name=Unc5b; Synonyms=Unc5h2;

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE

RP SPECIFICITY.

RX MEDLINE=97271897; PubMed=9126742;

RA Leonardo E.D., Hinck L., Masu M., Keino-Masu K., Ackerman S.L.,

RA Tessier-Lavigne M.;

RT "Vertebrate homologues of C. elegans UNC-5 are candidate netrin

RT receptors.";

RL Nature 386:833-838(1997).

RN [2]

RP FUNCTION, AND INTERACTION WITH DCC.

RX PubMed=10399920;

RA Hong K., Hinck L., Nishiyama M., Poo M.-M., Tessier-Lavigne M.,

RA Stein E.;

RT "A ligand-gated association between cytoplasmic domains of UNC5 and

RT DCC family receptors converts netrin-induced growth cone attraction to

RT repulsion.";

RL Cell 97:927-941(1999).

RN [3]

RP FUNCTION, AND MUTAGENESIS OF ASP-412.

RX PubMed=11387206; DOI=10.1093/emboj/20.11.2715;

RA Llambi F., Causeret F., Bloch-Gallego E., Mehlen P.;

RT "Netrin-1 acts as a survival factor via its receptors UNC5H and DCC.";

RL EMBO J. 20:2715-2722(2001).

CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates

CC axon repulsion of neuronal growth cones in the developing nervous

CC system upon ligand binding. Axon repulsion in growth cones may be

CC caused by its association with DCC that may trigger signaling for

CC repulsion. It also acts as a dependence receptor required for

CC apoptosis induction when not associated with netrin ligand.
 CC -!- SUBUNIT: Interacts with GNAI2 via its cytoplasmic part (By
 CC similarity). Interacts with the cytoplasmic part of DCC.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Mainly expressed in regions of differentiating
 CC neurons. Expressed in the developing sensory ganglia that flank
 CC the spinal cord from E12, peaking at E14. Expressed in the roof
 CC plate region of the spinal cord from E14.
 CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By
 CC similarity).
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
 CC cleavage does not take place when the receptor is associated with
 CC netrin ligand. Its cleavage by caspases is required to induce
 CC apoptosis.
 CC -!- SIMILARITY: Belongs to the UNC-5 family.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 ZU5 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U87306; AAB57679.1; -.
 DR HSSP; P07996; 1LSL.
 DR RGD; 621756; Unc5h2.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR011029; DEATH_like.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; Death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; TSP_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Apoptosis; Developmental protein; Immunoglobulin domain;
 KW Phosphorylation; Receptor; Repeat; Signal; Transmembrane.
 FT SIGNAL 1 26 Potential.
 FT CHAIN 27 945 Netrin receptor UNC5B.
 FT DOMAIN 27 377 Extracellular (Potential).
 FT TRANSMEM 378 398 Potential.

FT	DOMAIN	399	945	Cytoplasmic (Potential).
FT	DOMAIN	48	145	Ig-like.
FT	DOMAIN	153	242	Ig-like C2-type.
FT	DOMAIN	246	300	TSP type-1 1.
FT	DOMAIN	302	354	TSP type-1 2.
FT	DOMAIN	541	644	ZU5.
FT	DOMAIN	865	943	Death.
FT	SITE	412	413	Cleavage (by caspase-3).
FT	SITE	707	725	Interaction with DCC.
FT	DISULFID	69	128	By similarity.
FT	DISULFID	174	225	By similarity.
FT	CARBOHYD	222	222	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	347	347	N-linked (GlcNAc. . .) (Potential).
FT	MUTAGEN	412	412	D->N: Abolishes cleavage by caspase-3 and subsequent induction of apoptosis.
SQ	SEQUENCE	945 AA;	103520 MW;	6E9C2A262E560B9B CRC64;

Query Match 53.8%; Score 2578.5; DB 1; Length 945;
 Best Local Similarity 53.0%; Pred. No. 4.5e-185;
 Matches 509; Conservative 142; Mismatches 231; Indels 79; Gaps 17;

Qy	1	MAVRPGLWPALLGIVLAAW-----LRG--SGAQQSATVANPVPGANPDLLPHFLVEPEDV	53
		: : : :	
Db	1	MRARSGARGALLLALLLCWDPTPSLAGIDSGGQ---ALPDSFSPAPAEQLPHFLLEPEDA	57
Qy	54	YIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQV	113
		: : :	
Db	58	YIVKNKPVLEHCRAPATQIYFKCNGEWVSQKGHVTTQESLDEATGLRIREVQIEVSRQV	117
Qy	114	EKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPP	173
		: : : :	
Db	118	EELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPP	177
Qy	174	EGIPPAEVEWLRNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSA	233
		: : : : : : :	
Db	178	EGVPVAEVEWLNEDVIDPAQDTNFLTIDHNLIRQARLSDTANYTCVAKNIVAKRRST	237
Qy	234	SAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPLNGGAFCEGQNVQKTACAT	293
		: : : :	
Db	238	TATVIVYVNGGWSSWAIEWSPCSNRCGRGWQKRTRCTNPAPLNGGAFCEGQACQKTACTT	297
Qy	294	LCPVDGSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDTNRCTSDLCV---	350
		: : : : :	
Db	298	VCPVDGAWTEWSKWSACSTECAHWSRECMAPPQNGGRDCSGTLLDSKNCTDGLCVLNQ	357
Qy	351	HSASGPE-----DVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLSDVDADSS-IL	400
		: : : : : : :	
Db	358	RTLNDPKSRPLEPSGDVALYAGLVAVFVVLAVLMAVGVIYRRNCRDFDITDSSAAL	417
Qy	401	TSGFQPVSIKPSKADNPHELL--TIQPDLSSTTTTQYQSLCPRQDGSPKPFQLTNGHLLSP	458
		: : : : : : : :	
Db	418	TGGFHPVNFKTARPSNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLD	476
Qy	459	L-----GGG-----RHTLHHSSPTSEAEFEVS	480
Db	477	LPSLKIKVYDSSTIGSGAGLADGADLLGVLPPTYPGDFSRDTHFLHLRS-----A	527

Qy 481 RLSTQNYFRSLPRGTSNMTYGTFTNGLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHK 540
 | : | : | | | | : | | | | | | | | : | | : | | : | : |
 Db 528 SLGSQ-HLLGLPRDPSSSVSGTFGCLGGRLTIPGTGVSLLPNGAIPQGKFYDLYLRINK 586

Qy 541 PEDVRLPLA-GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCGS 599
 | | | : | | : | | | : | | | : | | | : | | | : | | | : |
 Db 587 TEST-LPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIFQLKTQAHQGH 645

Qy 600 WEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAARKRLKLLLFAPVAC 659
 | | : | : | | : | | | | : | : | : | | : | : | | : | |
 Db 646 WEEVVTLDDEETLNTPCYCQLEAKSCHILLDQLGTYVFTGESYSRSAVKRLQLAIFAPALC 705

Qy 660 TSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRSLHDPSSL 719
 | | | | : | | | | | | | | | | : | | | | : | | | | : | | :
 Db 706 TSLEYSRLVYCLEDTPAALKEVLELERTLGGYLVEEPKTLLEFKDSYHNLRSLHDIPHAH 765

Qy 720 WSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSIN 779
 | : | | | | | | | | | | : | | | | | | | | : | : | | | : | | :
 Db 766 WRSKLLAKYQEIPFYHVWNGSQKALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLH 825

Qy 780 FNITKDTRFAELLALESEAGVPAL--VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLA 837
 : : | | | | | | : | | | | | | | | : | | | | | | | |
 Db 826 TTLA-ETPAGSLDALCSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLA 884

Qy 838 QKLHLDSHLSFFASKPSPPTAMIILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 897
 | | | : | : | : | | | | | | : | | : | | : | : : : : : : :
 Db 885 QKLSMDRYLNYFATKASPTGVILDWEARQQDDGDLNSLASALEEMGKSEMLVAMTTDGD 944

Qy 898 C 898
 |
 Db 945 C 945

RESULT 11

UN5B_HUMAN

ID UN5B_HUMAN STANDARD; PRT; 945 AA.
 AC Q8IZJ1; Q86SN3; Q8N1Y2; Q9H9F3;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Netrin receptor UNC5B precursor (Unc-5 homolog B) (Unc-5 homolog 2)
 DE (p53-regulated receptor for death and life protein 1)
 DE (UNQ1883/PRO4326).
 GN Name=UNC5B; Synonyms=P53RDL1, UNC5H2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2), TISSUE SPECIFICITY, AND INTERACTION
 RP WITH GNAI2.
 RC TISSUE=Lung;
 RX MEDLINE=22246081; PubMed=12359238; DOI=10.1016/S0006-291X(02)02277-5;
 RA Komatsuzaki K., Dalvin S., Kinane T.B.;
 RT "Modulation of G(ialpha(2)) signaling by the axonal guidance molecule
 RT UNC5H2.";
 RL Biochem. Biophys. Res. Commun. 297:898-905(2002).

RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, AND MUTAGENESIS OF ASP-412.
 RX PubMed=12598906; DOI=10.1038/ncb943;
 RA Tanikawa C., Matsuda K., Fukuda S., Nakamura Y., Arakawa H.;
 RT "p53RDL1 regulates of p53-dependent apoptosis.";
 RL Nat. Cell Biol. 5:216-223(2003).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
 RA Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.-H., Yansura D.,
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
 RA Godowski P., Gray A.;
 RT "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genome Res. 13:2265-2270(2003).
 RN [4]
 RP SEQUENCE OF 361-945 FROM N.A.
 RC TISSUE=Amygdala, and Teratocarcinoma;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).

RN [5]
 RP DOWN-REGULATION IN CANCER.
 RX PubMed=12655055; DOI=10.1073/pnas.0738063100;
 RA Thiebault K., Mazelin L., Pays L., Llambi F., Joly M.-O.,
 RA Scoazec J.-Y., Saurin J.-C., Romeo G., Mehlen P.;
 RT "The netrin-1 receptors UNC5H are putative tumor suppressors
 RT controlling cell death commitment.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:4173-4178(2003).
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates
 CC axon repulsion of neuronal growth cones in the developing nervous
 CC system upon ligand binding. Axon repulsion in growth cones may be
 CC caused by its association with DCC that may trigger signaling for
 CC repulsion. It also acts as a dependence receptor required for
 CC apoptosis induction when not associated with netrin ligand.
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC (By
 CC similarity). Interacts with GNAI2 via its cytoplasmic part.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8IZJ1-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8IZJ1-2; Sequence=VSP_011698;
 CC -!- TISSUE SPECIFICITY: Highly expressed in brain. Also expressed at
 CC lower level in developing lung, cartilage, kidney and
 CC hematopoietic and immune tissues.
 CC -!- INDUCTION: By p53/TP53.
 CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By
 CC similarity).
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
 CC cleavage does not take place when the receptor is associated with
 CC netrin ligand. Its cleavage by caspases is required to induce
 CC apoptosis.
 CC -!- MISCELLANEOUS: Down-regulated in multiple cancers including
 CC colorectal, breast, ovary, uterus, stomach, lung, or kidney
 CC cancers.
 CC -!- SIMILARITY: Belongs to the UNC-5 family.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 ZU5 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AY126437; AAM95701.1; -.
 DR EMBL; AB096256; BAC57998.1; -.
 DR EMBL; AY358351; AAQ88717.1; -.
 DR EMBL; AK022859; BAB14276.1; ALT_INIT.
 DR EMBL; AK094595; BAC04382.1; ALT_INIT.
 DR HSSP; P07996; 1LSL.

DR Genew; HGNC:12568; UNC5B.
 DR MIM; 607870; -.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR011029; DEATH_like.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; Death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; TSP_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Alternative splicing; Apoptosis; Developmental protein;
 KW Immunoglobulin domain; Phosphorylation; Polymorphism; Receptor;
 KW Repeat; Signal; Transmembrane.
 FT SIGNAL 1 26 Potential.
 FT CHAIN 27 945 Netrin receptor UNC5B.
 FT DOMAIN 27 377 Extracellular (Potential).
 FT TRANSMEM 378 398 Potential.
 FT DOMAIN 399 945 Cytoplasmic (Potential).
 FT DOMAIN 48 145 Ig-like.
 FT DOMAIN 147 242 Ig-like C2-type.
 FT DOMAIN 246 300 TSP type-1 1.
 FT DOMAIN 302 354 TSP type-1 2.
 FT DOMAIN 541 644 ZU5.
 FT DOMAIN 865 943 Death.
 FT SITE 412 413 Cleavage (by caspase-3).
 FT SITE 707 725 Interaction with DCC (By similarity).
 FT DISULFID 69 128 By similarity.
 FT DISULFID 174 225 By similarity.
 FT CARBOHYD 222 222 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 347 347 N-linked (GlcNAc. . .) (Potential).
 FT VARSPLIC 356 367 NKKTLSDPNSHL -> M (in isoform 2).
 FT /FTId=VSP_011698.
 FT VARIANT 516 516 A -> T (in dbSNP:10509332).
 FT /FTId=VAR_019730.
 FT MUTAGEN 412 412 D->N: Abolishes cleavage by caspase-3 and
 FT subsequent induction of apoptosis.
 FT CONFLICT 483 483 K -> E (in Ref. 3).
 FT CONFLICT 851 851 L -> P (in Ref. 3; BAB14276).
 SQ SEQUENCE 945 AA; 103637 MW; 56064E335F323447 CRC64;

Query Match 53.4%; Score 2558.5; DB 1; Length 945;
 Best Local Similarity 52.7%; Pred. No. 1.5e-183;
 Matches 501; Conservative 148; Mismatches 244; Indels 57; Gaps 15;

Qy 1 MAVRPGLWPALLGIVLAAW---LRGSGAQQ-SATVANPVPGANPDLLPHFLVEPEDVYIV 56
 | | | ||| :| | | :| | : : | | : ||:| ||:| |||

Db 1 MGARSGARGALLLALLLCWDPRLSQAGTDSGSEVLPSDFPSAPAEPLPYFLQEPQDAYIV 60

Qy 57 KNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQVEKV 116
| | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 61 KNKPVELRCRAFPATQIFYKCNGEWVSQNDHVTQEGLEATGLRVREVQIEVSRQQVEEL 120

Qy 117 FGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGI 176
| | | : | | | | | | : | | | | : | | | | : | | | | : | | | | :

Db 121 FGLEDYWCQCVAWSSAGTTKSRRAYVRIAYLRKNFDQEPLGKEVPLDHEVLLQCRPPEGV 180

Qy 177 PPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAA 236
| | | | | : | | : | | : | | : | | : | | : | | : | | : | | :

Db 181 PVAEVEWLKNEDVIDPTQDTNFLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRSTTAT 240

Qy 237 VIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCP 296
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 241 VIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRCTNPAPLNGGAFCEGQAFQKTACTTICP 300

Qy 297 VDGSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHS---- 352
| | : | : | | | | : | | | | | | : | | | : | | | : | | : :

Db 301 VDGAWTEWSKWSACSTECAHWSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCMQNKKTL 360

Qy 353 -----ASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLSDSDVADSS-ILT 401
| | | | | | | | : | : | : | : | : | : | : | : | : | : | :

Db 361 SDPNSHLLEASG--DAALYAGLVVAIFVVVAILMAVGVVVYRRNCRDFDTDITDSSAALT 418

Qy 402 SGFQPVSIKPSKADNPHLL--TIQPDLSSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPL 459
| | | : | : | | | : | | : | : | : | : | : | : | : | : | :

Db 419 GGFHPVNFKTARPSNPQLLHPSVPPDLTASAGIYRGPVYALQDS-TDKIPMTNSPLLDPL 477

Qy 460 GGGRHTLHHSSPT-----SEAEFVSRLSTQNY-----FRS-----L 491
: : | | | : : | : | : | : | : | : | : | : | : | : | :

Db 478 PSLKVKVYSSSTTGSGPGLADGADLLGVLP PGTYPSDFARDTHFLHLRSASLGSQQLLGL 537

Qy 492 PRGTSNMTYGTENFLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLA-G 550
| | : | | | | | | | : | | : | | | : | | : | | : | | : | :

Db 538 PRDPGSSVSGTFGCLGGRLSIPGTGVSLLPNGAIPQGKFYEMYLLINKAEST-LPLSEG 596

Qy 551 CQTLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCGWSWEDVLHLGEEA 610
| | : | | | : | | | : | | | | | | : | | | : | | : | | | :

Db 597 TQTVLSPSVTCGPTGLLLCRPVILTMPHCAEVSARDWIFQLKTQAHQGHWEVVTLDEET 656

Qy 611 PSHLYYCQLEASACYVFTEQLGRFALVGEALSVAALKRLKLLLFAPVACTSLEYNIRVYC 670
: | | | | | : | : | : | : | : | : | : | : | : | : | : | :

Db 657 LNTPCYCQLEPRACHILLDQLGTYVFTGESYSRSAVKRLQLAVFAPALCTSLEYSRLVYC 716

Qy 671 LHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKSLLVSYQE 730
| | | | | : | : | | | : | : | : | | | | | | : | : | | | :

Db 717 LEDTPVALKEVLELERTLGGYLV EEPKPLMFKDSYHNLRSLHDLPHAHWRSKLLAKYQE 776

Qy 731 IPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAE 790
| | | | | : | : | | | | | : : : | | : | | | : | | : : | :

Db 777 IPFYHIWGSQKALHCTFTLERHSLASTELTCKICVRQVEGEGQIFQLHTTLA-ETPAGS 835

Qy 791 LLALESEAG--VPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSF 848
| | | | | : | | | | | | | : | | | | | | | : | : : :

Db 836 LDTLCSAPGSTVTTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRMLAQKLSMDRYLNY 895

Qy 849 FASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898
 ||:| ||| :||:|||| :|:|: ||:|: :|: : : : : :|
 Db 896 FATKASPTGVILDLWEALQQDDGDLNSLASALEEMGKSEMLVAVATDGDC 945

RESULT 12

UN5D_MOUSE

ID UN5D_MOUSE STANDARD; PRT; 956 AA.
 AC Q8K1S2;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Netrin receptor UNC5D precursor (Unc-5 homolog D) (Unc-5 homolog 4).
 GN Name=Unc5d; Synonyms=Unc5h4;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=22239710; PubMed=12351186; DOI=10.1016/S0925-4773(02)00248-4;
 RA Engelkamp D.;
 RT "Cloning of three mouse Unc5 genes and their expression patterns at
 RT mid-gestation.";
 RL Mech. Dev. 118:191-197(2002).
 CC -!- FUNCTION: Receptor for netrin involved in cell migration. May be
 CC involved in axon guidance by mediating axon repulsion of neuronal
 CC growth cones in the developing nervous system upon ligand binding.
 CC Axon repulsion in growth cones may be caused by its association
 CC with DCC that may trigger signaling for repulsion. It also acts as
 CC a dependence receptor required for apoptosis induction when not
 CC associated with netrin ligand (By similarity).
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in developing limb and mammary
 CC gland.
 CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By
 CC similarity).
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
 CC cleavage does not take place when the receptor is associated with
 CC netrin ligand. Its cleavage by caspases is required to induce
 CC apoptosis (By similarity).
 CC -!- SIMILARITY: Belongs to the UNC-5 family.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 ZU5 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AJ487854; CAD32252.1; -.
DR HSSP; P07996; 1LSL.
DR MGD; MGI:2389364; Unc5d.
DR InterPro; IPR000488; Death.
DR InterPro; IPR011029; DEATH_like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00531; Death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00209; TSP1; 2.
DR PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50092; TSP1; 2.
KW Apoptosis; Developmental protein; Immunoglobulin domain;
KW Phosphorylation; Receptor; Repeat; Signal; Transmembrane.
FT SIGNAL 1 30 Potential.
FT CHAIN 31 956 Netrin receptor UNC5D.
FT DOMAIN 31 382 Extracellular (Potential).
FT TRANSMEM 383 403 Potential.
FT DOMAIN 404 956 Cytoplasmic (Potential).
FT DOMAIN 52 149 Ig-like.
FT DOMAIN 151 242 Ig-like C2-type.
FT DOMAIN 250 304 TSP type-1 1.
FT DOMAIN 306 358 TSP type-1 2.
FT DOMAIN 543 645 ZU5.
FT DOMAIN 862 939 Death.
FT SITE 419 420 Cleavage (by caspase-3) (By similarity).
FT SITE 706 724 Interaction with DCC (By similarity).
FT DISULFID 73 132 By similarity.
FT DISULFID 178 229 By similarity.
FT CARBOHYD 115 115 N-linked (GlcNAc . . .) (Potential).
FT CARBOHYD 226 226 N-linked (GlcNAc . . .) (Potential).
FT CARBOHYD 351 351 N-linked (GlcNAc . . .) (Potential).
FT CARBOHYD 379 379 N-linked (GlcNAc . . .) (Potential).
SQ SEQUENCE 956 AA; 106351 MW; DFDF07839C10C68D CRC64;

Query Match 45.9%; Score 2200; DB 1; Length 956;

Best Local Similarity 45.5%; Pred. No. 1.5e-156;

Matches 431; Conservative 159; Mismatches 280; Indels 78; Gaps 15;

Qy 8 WPALLGIVLAAWLRGS----GAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLL 63
| | : | : | | : : : | | | | : | : | : |
Db 15 WLPWLGLFF--WAAGAAAARGADGSEILPDSIPSA-PGTLPHFIEEPEDAYIIKSNPIAL 71

Qy 64 VCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGLLEEW 123
| | | | | | | | | | : | | | | | | | | | : | : | : |
Db 72 RCKARPAMQIFFKCNGEWVHQNEHVSEESLDESSGLKRVREVFINVTRQQVEDFHGPEDYW 131

Qy	124	CQCVAWSSSGTTKSKAYIRIARLRKNFEQEPLAKEVSLSEQGIVLPCRPEGIPPAEVEW	183
Db	132	:: : : : : : :	
		CQCVAWSHLGTSKSRKASVRIAYLRKNFEQDPQGREGVPIEGMIVLHCRPPEGVPAAEVEW	191
Qy	184	LRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASA AAVIVYVNG	243
Db	192	:: : : : : ::: : : : : :	
		LKNEEPIDSEQDENIDTRADHNLIIRQARLSDSGNYTCMAANIVAKRRSLSATVVVYVNG	251
Qy	244	GWSTWTEWSVCSASCGRGWQKR SR SCTNPAPLN GGAFCEGQN VQKTACATLC PVDGSWSP	303
Db	252	: : : : : : :	
		GWSSWTEWSACNVRCGRGWQKR SRTCTNPAPLN GGAFCEGMSVQKITCTALCPVDGSWEV	311
Qy	304	WSKWSACGLDCTHWRSRECS DPAPR NGGE ECQGTDL DTRNCTSDLCV-----	350
Db	312	: : : : : :: :	
		WSEWSVCSPECEHLRI RECTAPP PRNG GKFC EGLSQESENCTDGLCILDKKPLHEIKPQR	371
Qy	351	---HSASGPEDVALYVGLIAVAVCLVLLLVLILVYCRKKEGLSDSDVADSSILTSGFQPV	407
Db	372	: : : : :	
		WSRRGIENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSHSDYGVDVIDSSALTGGGFQTF	430
Qy	408	SIKPSKADNPHELL--TIQPD LSTTTTTTYQGS LCPRQDG PSPKFQLTN GHLLSPLG----	460
Db	431	: : : : : : :	
		NFKTVRQGN SLLLPAM QPDL-TVS RTYSGPIC-LQD- PLDKELMTES SLFNPLSDIKVK	487
Qy	461	-----GGRH-----TLHHSSPTSEAE EFVSRLSTQ NYFR	489
Db	488	: : ::	
		VQSSEFMVSLGV SERAEYHGKNHSGTFPHGNNR GFSTIHPRNK T----PYIQNL S-----	537
Qy	490	SLPRGTSNM TYGT FNFLGGRL MIPNTGIS LLIPPDA IPRGKIYEI YLT LHKPEDV RLPLA	549
Db	538	: : : :::	
		SLPTRTELRTTG VFHGLGGRLVM PNTGVS LLIPHGA IPEENS WEIYMSIN QGEP-SLQSD	596
Qy	550	GCQTLLSPIVSCGPPGVLLTRPVILAMDHC GEPS PD SWSLRLKKQ SCEGS WEDVLHLGEE	609
Db	597	: : : : : : : : : : : : :	
		GSEVLLSPEVTC GPPDMLVT TPFA LTIPHCADV SSEHWN IHLKKRTQQG KWEEVMS VEDE	656
Qy	610	APSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLL FAPVACTSLEY NIRVY	669
Db	657	: : : : : :: : : : : : :	
		STS--CYCLLDPFACHVLLDSFGTYALTGE PITDC AVKQLKVAVFGCM SCNSLDYNLRVY	714
Qy	670	CLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF KD SYHN LR LS IHDPSSLWK SKLLVS YQ	729
Db	715	: : : : : : : : : : : : : : :	
		CVDNTPCAFAQEVISDERHQGGQLLEEPKLLHFKGNTFSLQVSVLDIPFFLWR IKPFTACQ	774
Qy	730	EIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFA	789
Db	775	: : :: : : : : : : : : : :	
		EVPFSRVWSSNRQPLHCAFSLERYTPTTTQLSCKICIRQLKGHEQILQVQTSILESERET	834
Qy	790	ELLALESEAGVPALVGPSAFKIPFLIRQKI ISSLDPPCRRGADWRTL AQKLHLD SHLSFF	849
Db	835	: :: : : :: : : :: : :	
		ITFFAQEDSTFP AQ TGPKAFKIPYSIRQR ICATFDTPNA KGKDWQMLAQ KNSINRNLSYF	894
Qy	850	ASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	897
Db	895	:: : : : : : : :	
		ATQSSPSAVILNLWEARHOODGDLDSLACALEEIGRTHTKLSNITEPO	942

RESULT 13

UN5D_HUMAN

ID UN5D_HUMAN STANDARD; PRT; 953 AA.
AC Q6UXZ4; Q8WYP7;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Netrin receptor UNC5D precursor (Unc-5 homolog D) (Unc-5 homolog 4)
DE (UNQ6012/PRO34692).
GN Name=UNC5D; Synonyms=KIAA1777, UNC5H4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.-H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P., Gray A.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RA Nakajima D., Nakayama M., Nagase T., Ohara O.;
RT "Identification of unc5H4 gene.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Receptor for netrin. May be involved in axon guidance by
CC mediating axon repulsion of neuronal growth cones in the
CC developing nervous system upon ligand binding. Axon repulsion in
CC growth cones may be caused by its association with DCC that may
CC trigger signaling for repulsion. It also acts as a dependence
CC receptor required for apoptosis induction when not associated with
CC netrin ligand (By similarity).
CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q6UXZ4-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q6UXZ4-2; Sequence=VSP_011703;
CC Note=No experimental confirmation available;
CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By
CC similarity).
CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The

CC cleavage does not take place when the receptor is associated with
 CC netrin ligand. Its cleavage by caspases is required to induce
 CC apoptosis (By similarity).
 CC -!- SIMILARITY: Belongs to the UNC-5 family.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 ZU5 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB055056; BAB83663.1; -.
 DR EMBL; AY358147; AAQ88514.1; -.
 DR HSSP; P07996; 1LSL.
 DR Genew; HGNC:18634; UNC5D.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR011029; DEATH_like.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; Death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; TSP_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Alternative splicing; Apoptosis; Developmental protein;
 KW Immunoglobulin domain; Phosphorylation; Receptor; Repeat; Signal;
 KW Transmembrane.
 FT SIGNAL 1 32 Potential.
 FT CHAIN 33 953 Netrin receptor UNC5D.
 FT DOMAIN 33 379 Extracellular (Potential).
 FT TRANSMEM 380 400 Potential.
 FT DOMAIN 401 953 Cytoplasmic (Potential).
 FT DOMAIN 54 151 Ig-like.
 FT DOMAIN 153 244 Ig-like C2-type.
 FT DOMAIN 252 306 TSP type-1 1.
 FT DOMAIN 308 360 TSP type-1 2.
 FT DOMAIN 540 642 ZU5.
 FT DOMAIN 859 936 Death.
 FT SITE 416 417 Cleavage (by caspase-3) (By similarity).
 FT SITE 703 721 Interaction with DCC (By similarity).
 FT DISULFID 75 134 By similarity.


```
DR    PROSITE; PS50092; TSP1; 2.
FT    NON_TER          1          1
FT    NON_TER          876        876
SQ    SEQUENCE          876 AA;   96301 MW;  6F9336D53E096E00 CRC64;
```

Qy	43	LPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM	102
Db	2	LPVFLVEPKGAYVMKNRPALKYCKASHALQISFKCSGS-TKPPPTKEKHHTDPHSGVQLQ	60
Qy	103	EVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSL	162
Db	61	EATATITRELVDDEFFGKGPFKCECRAYSSRGHVKTQPVTIQVATIKKQISISPKIVRVAT	120
Qy	163	EQGIVLPCRPPEGIPPAEVEWLRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCV	222
Db	121	GGRAELNC-IANATPAKVWLKNS--VPVHANPPFVLLTENALLIARVEIQDMANYTCV	177
Qy	223	AKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSA-SCGRGWQKRSRSCTNPAPLNGGAF	281
Db	178	AENIAGKRVSDFVPITVYVDGGWSSWGPWTDCKCPGHGKQGQKRTRVCNSPVPMMNGAP	237
Qy	282	EGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWRSRSCSDPA-PRNGGE--ECQGT	338
Db	238	KGASTESTPDCLPCSA-GRWSSWSEWSECGPDCTQIRQRSCVAQAFAIDSGTIVNCAGKS	296
Qy	339	LDTRNCTSDLCVHSASGPE-DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADS	397
Db	297	QOSIKCTGGLCNYTAQDSNWSVYLWVTLVA-AFCLGVVFAV--SKFLRRKKTIPAYNLAR	353
Qy	398	SILTSGFQPVSIK-PSKAD-NPHLLTIQP-----DLSTTTTT	432
Db	354	SDLTQNTGPIINYEYPMASLQPHLAGHHPHHHHLQQQQHHHHLSSLHHQHGGSSLLGPVAT	413
Qy	433	YQGS---CPRQDGPSPKFQLTNGHLLSPLGGGRHTL-----HHSS-----	470
Db	414	GHGQLHPQCQSQAAPT-----LPIGGLKSSLPLPRSNSEHHYDVPHLNNTSPA	462
Qy	471	--PTSEAEFVSRLSTQNYFRSLPRGTSNMITYGTNFGGLMIPNTGISLLIPDAIPR	528
Db	463	IRPTTLVEE-----PFRGAEVTHATLTAPAGALLRLATYSTALLIPEGAIPK	508
Qy	529	GKIYEIYLTLLHKPE--DVRPLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPSD	586
Db	509	HQRHSVALSIVRDDKHHPVPPTGPRSTYLSPPVFCGPVDTKVNKPIVMQLPHCAENLSD	567
Qy	587	WSRLRLKKQSCGS-WEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS---	642
Db	568	WAFSLYSAPDNVTPWCKVVTIGEETLNTPALVQIDKRYAYVLTETFGKYVLVGESATDIQ	627
Qy	643	-VAAAKRLKLLLFPVACTSL-EYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLH	700
Db	628	ERVACKRLRLFICGPSTVPEFSDVSLRVYIVEDNPGAEECRHCEQEIGGVLLGRSTVLH	687

Qy 701 FKD-SYHNLRLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSD 759
 | | | : : | : : | | | : | | : | |
 Db 688 FADVGQAGLNIDLQCVGG--WRAKSSSERQTIPFSHVWNSACTALHCSFTL----- 736

Qy 760 LACKLWVWQVEGDGQSFSINFNITKDTIFA---ELLALESEAGVPA----- 802
 | : | | | : : | | : | | |
 Db 737 --CR-----TEHDKCDFKIVVQASQDVPQGLDERLTAI-----GVPATLSISSVSGSDHNT 785

Qy 803 -LVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDHSHLSFFASKPSPTAMILN 861
 | | | : : : : | | | : | | | : : | : : | : : | : | : | :
 Db 786 NLVATDRFRLSKDVKRKLCRLDPPTQKRNDWRMLAAHLNVDRYLTYFATRPSPPTDQILD 845

Qy 862 LWEARHFPNGNLSQLAAVAGLGQPD 888
 | | | | : | | : : | |
 Db 846 LWEARNRDLNALQQLIEICRTMERPD 872

RESULT 15

UNC5_DROME

ID UNC5_DROME STANDARD; PRT; 1072 AA.
 AC Q95TU8; Q9NBL0; Q9V7B5;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-JAN-2005 (Rel. 46, Last annotation update)
 DE Netrin receptor unc-5 precursor (Unc5 netrin receptor).
 GN Name=unc-5; ORFNames=CG8166;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RX PubMed=11719202;
 RA Keleman K., Dickson B.J.;
 RT "Short- and long-range repulsion by the Drosophila Unc5 netrin
 RT receptor."
 RL Neuron 32:605-617(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP GENOME REANNOTATION.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley; TISSUE=Embryo;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celniker S.E.;
 RT "A *Drosophila* full-length cDNA resource.";
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 RN [5]
 RP FUNCTION.
 RX PubMed=12765609;
 RA Freeman M.R., Delrow J., Kim J., Johnson E., Doe C.Q.;
 RT "Unwrapping glial biology: Gcm target genes regulating glial
 RT development, diversification, and function.";
 RL Neuron 38:567-580(2003).
 CC -!- FUNCTION: Receptor for netrin required for motor axon guidance.
 CC Mediates both short- and long-range axon motor repulsion in the
 CC developing nervous system upon ligand binding. Also involved in
 CC glial migration. While short-range repulsion requires both fra and
 CC unc-5, long-range repulsion only requires unc-5.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- TISSUE SPECIFICITY: Prior to gastrulation, it is strongly
CC expressed in the presumptive mesoderm. Mesodermal expression
CC begins to fade during stages 13-14, persisting only in the cells
CC that form the dorsal vessel. Expressed within the CNS from late
CC stage 13, shortly after the first axons have extended. Detected in
CC several dispersed clusters of cells within the CNS, increasing in
CC number as development proceeds. Also expressed in the peripheral
CC and exit glia, which migrate laterally out of the CNS between
CC stages 14 and 17. Strongly expressed in motor axons that exit the
CC CNS ipsilaterally via the segmental nerve root (SN). Not expressed
CC on either commissural or longitudinal axons within the CNS, nor on
CC motor axons that exit via the intersegmental nerve (ISN). In the
CC periphery, it is detected on all branches of the SN. Also
CC expressed at high level in exit and peripheral glia along both the
CC SN and ISN.
CC -!- PTM: Phosphorylated on different cytoplasmic tyrosine residues (By
CC similarity).
CC -!- SIMILARITY: Belongs to the UNC-5 family.
CC -!- SIMILARITY: Contains 1 death domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; AF247762; AAF74193.1; -.
DR EMBL; AE003811; AAF58143.2; -.
DR EMBL; AY058501; AAL13730.1; -.
DR HSSP; P07996; 1LSL.
DR FlyBase; FBgn0034013; unc-5.
DR GO; GO:0007411; P:axon guidance; IGI.
DR GO; GO:0008347; P:glia cell migration; IMP.
DR InterPro; IPR000488; Death.
DR InterPro; IPR011029; DEATH_like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00531; Death; 1.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00209; TSP1; 2.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50092; TSP1; 2.

KW Developmental protein; Immunoglobulin domain; Phosphorylation;
KW Receptor; Repeat; Signal; Transmembrane.

FT SIGNAL 1 30 Potential.
FT CHAIN 31 1072 Netrin receptor unc-5.
FT DOMAIN 31 440 Extracellular (Potential).

FT	TRANSMEM	441	461	Potential.
FT	DOMAIN	462	1072	Cytoplasmic (Potential).
FT	DOMAIN	128	224	Ig-like.
FT	DOMAIN	232	314	Ig-like C2-type.
FT	DOMAIN	324	379	TSP type-1 1.
FT	DOMAIN	398	499	TSP type-1 2.
FT	DOMAIN	980	1067	Death.
FT	DISULFID	149	207	By similarity.
FT	DISULFID	253	303	By similarity.
FT	CARBOHYD	79	79	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	300	300	N-linked (GlcNAc. . .) (Potential).
FT	CONFLICT	885	885	P -> S (in Ref. 1).
SQ	SEQUENCE	1072 AA;	116416 MW;	A0A6B5A96B10138F CRC64;

Query Match 20.7%; Score 992; DB 1; Length 1072;
 Best Local Similarity 28.7%; Pred. No. 1.7e-65;
 Matches 282; Conservative 145; Mismatches 383; Indels 174; Gaps 33;

Qy	36	PGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDG	95
		: : : : : :	
Db	120	PGEASNTLPIFLIEPESVFVVKNRPAVLKCKASHSLQVIFKCSGS-SQPPPSTHETHVDP	178
Qy	96	SSGLPTMEVRINVSRRQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP	155
		: : : : : : : :	
Db	179	HTGVNMEEVTATIHRDLVDEFFGDGPFKCECHAWSSRGVVKSAATVHIAYIRKSFNQSP	238
Qy	156	LAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDPNVYITREHSLVVRQARLAD	215
		: : : : : : : : : : : :	
Db	239	TSLRLELGSRAELRCEPPGGFPEPKLTWHKNNNAVITADSEPGITVS-AGTLIFRQVALQH	297
Qy	216	TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSA-----SCGRGWQKRSRSC	271
		: : : :	
Db	298	MANYSCSAENIAGRRVSDSAVLIVYVNGGWSTWSPWRECKCAGKPSQGR---KRSRTCNN	354
Qy	272	PAPLNGGAFCEGQNVQKTACATLCPVD-----GSWSPWSKWSACGLDC	314
		: :	
Db	355	PMPLNGGAQCPCGPQIQKSADCAACPEDTQIVSPDGFDISSSKRMARWSAWSWDSICSAEC	414
Qy	315	THWRSREC-----SDP	325
		:	
Db	415	IQVRRRKCLTQGGTQISSEAEAGDLLLGAPGVGMAALIAAGVGAVGSPSEATGSSSDI	474
Qy	326	APRNGGEECQGTDLDRNCTSDLCVHSASGPEDV--ALYVGLIAV-AVCLVLLLLVLILV	382
		: : : :	
Db	475	IPGYGKSLCAGKDIQTAECRGEQC---QIGKDDFDWTLYLGLAFITAVCFAGFTALI---	528
Qy	383	YCRKKEGLDSDV---ADSSILTSQFQPVSIKPSKADNPHLLTIQPDLSSTTTTQYQSLCP	439
		: : : : : : : :	
Db	529	-CCARRGIRTNPHYNMARSVMDADYMP-GVVEKKEMRMHI-----EASNMGYDYV-----	576
Qy	440	RQDGPSPKQFQLTNGHLLS-PLGGGRHTLHH-----SSPTSEAE	477
		: : :	
Db	577	---NPGHRY-LPGEHIMGMGIGCGGVTEHHYDVPNLSANYTNPIDHLSVDYLSETGESST	632
Qy	478	FVSRLST--QNYFRSLPRGTSNMTYGTNFGRLMIPNTGISLLIPPDAIPRGKIYEIY	535
		: : : : : : :	
Db	633	ADTSNSTFDMNGKLSILNASKSSTYEMLGSAAGQLRLYGGELLFLFVPEHAIGHVKKHVS	692

Qy 536 LTLHKPEDVRLPLAGCQT----LLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRL 591
 | | | | : | | | : | | : | | | : | | : | : | : | :
 Db 693 LLLLSDECSRV---SCATESSILCSSVVHSA PRNYSFVKPVILKIPHC-LVAPEQWHVHI 748

Qy 592 KKQSCEG-----SWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 646
 | : | : : | | : : | | | : : | | | | : | | :
 Db 749 YHADSEHDELSVNWRRRAVSVEETINTPMFVQLEATHVFIMTEQLGHFTVVAEPRIQQPS 808

Qy 647 KRLKLLLFAP-VACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSY 705
 : : | | | : : : : | : : : : | : | | : | : | :
 Db 809 IKMKLLAFSQHTPPSNANCSLRIYVVKDFPN SRDICANVEAKLGG SFLGESQVFAFTLNS 868

Qy 706 HNLRLSIH--DVPSSLWKSKLLVSYQE-IPFYHIWNGTQRYLHCTFTLERVSPSTSDLAC 762
 | | : : | | : : | : | | : : | | | : | : : |
 Db 869 RNLNIRVRSADVEAA-----APYEHAIPYQHILSNNS-ILHCEFSLRRQDQNS---LC 917

Qy 763 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPA-----LVGPSA-----FKIPF 813
 : | | : : | : : | | | : : | | : : |
 Db 918 VDFGQGS EDDYYTFNIPAHSMGS--AEELASTTNTTISIDRQGN YVNESCVMD FVQLPH 975

Qy 814 LIRQKIISLDPPCRRGADWRTL AQKLHLD SHLSFFASKPSPTAMILNLWEAR--HFPNG 871
 : : | : | | | | | : : | : : | : : | : | | | | | |
 Db 976 ATKRLICGALDPPRADERDWRL LAKKLNTDRYIAYFATKASPT EQILNLWECRANSSPGS 1035

Qy 872 NLSQLAAAVAGLGQPDAGLFTVSE 895
 : : : : : | | : |
 Db 1036 SSNSVSHTIMAL-----LLTLKE 1053

Search completed: March 1, 2005, 09:03:30
 Job time : 169.273 secs